



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

МИЭМ

МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОНИКИ
И МАТЕМАТИКИ им. А.Н.ТИХОНОВА

МАТЕРИАЛЫ КОНФЕРЕНЦИИ

2023

МЕЖВУЗОВСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ

КОНФЕРЕНЦИЯ

СТУДЕНТОВ, АСПИРАНТОВ
И МОЛОДЫХ СПЕЦИАЛИСТОВ

им. Е.В.АРМЕНСКОГО



**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**МОСКОВСКИЙ ИНСТИТУТ ЭЛЕКТРОНИКИ И МАТЕМАТИКИ им.А.Н.Тихонова
НАЦИОНАЛЬНОГО ИССЛЕДОВАТЕЛЬСКОГО УНИВЕРСИТЕТА
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**Межвузовская научно-техническая
конференция студентов, аспирантов
и молодых специалистов
имени Е.В. Арменского**

МАТЕРИАЛЫ КОНФЕРЕНЦИИ

Москва 2023 г.

ББК 2+3
Н 34

Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов имени Е.В. Арменского. Материалы конференции. - М. ~: МИЭМ НИУ ВШЭ, 2023. – 314 стр.

ISBN 978-5-903650-75-0

В материалах конференции студентов, аспирантов и молодых специалистов представлены тезисы докладов по следующим направлениям: математика и компьютерное моделирование; информационно-коммуникационные технологии; автоматизация проектирования, банки данных и знаний, интеллектуальные системы; компьютерные образовательные продукты; информационная безопасность; электроника и приборостроение; производственные технологии, нанотехнологии и новые материалы.

Материалы конференции могут быть полезны для преподавателей, студентов, научных сотрудников и специалистов, специализирующихся в области прикладной математики, информационно-коммуникационных технологий, электроники и информационной безопасности.

Редакционная коллегия: Е.А. Крук, С.А. Аксенов, Г.Г. Бондаренко, Л.С. Восков,
А.А. Елизаров, Ф.И. Иванов, А.Б. Лось, Н.С. Титкова

Издание осуществлено с авторских оригиналов.

ISBN 978-5-903650-75-0

© Московский институт электроники и
математики Национального
исследовательского университета
«Высшая школа экономики», 2023 г.
© Авторы, 2023 г.

Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В. Армевского названа в честь основателя и первого ректора Московского института электроники и математики. Евгений Викторович в 1962 году создал вуз на стыке самых передовых образовательных методик того времени в области электроники и реального сектора экономики, представленного крупнейшими предприятиями в этой области.

Краткая биографическая справка: Евгений Викторович Армений родился 2 октября 1923 года в городе Буй Костромской области. В 1941 году он ушел добровольцем на фронт в составе 234-й Ярославской стрелковой коммунистической дивизии. Первый фронтовой опыт получил в боях под Москвой. В самые тяжелые 1941-1942 годы он был разведчиком-наблюдателем 1081 артполка на Калининском фронте, а после ранения в конце 1942 года - курсантом учебной роты связи 5-ой отдельной запасной стрелковой бригады Уральского военного округа. С 1943 года по ноябрь 1945 года был командиром радиовзвода 102 стрелкового корпуса Первого Украинского фронта. Великую Отечественную войну Евгений Викторович закончил в Праге 9 мая 1945 года. За ратные подвиги он был награжден орденами Красной Звезды, Отечественной войны I и II степени, медалями «За боевые заслуги», «За освобождение Праги», «За победу над Германией в Великой Отечественной войне».



После демобилизации он поступил в Ивановский химико-технологический институт, в котором проучился 2 курса, а затем перешел на третий курс Московского механического института (ныне Московский инженерно-физический институт, НИЯУ МИФИ), который закончил в 1951 году.

Трудовую деятельность Евгений Викторович начал с должности секретаря комитета ВЛКСМ МИФИ, а с 1952 года началась его педагогическая деятельность ассистентом кафедры «Автоматика и телемеханика» МИФИ. В этом же году он поступил в аспирантуру МИФИ, которую успешно завершил, и защитил кандидатскую диссертацию на тему «Пространственный метод регулирования двухфазного асинхронного двигателя». С 1956 года Евгений Викторович развивает научное направление и формирует научную школу в области создания систем автоматического контроля, регулирования и управления электрофизических установок.

12 апреля 1962 года Советское правительство приняло решение о создании Московского института электронного машиностроения. Евгений Викторович Армений был назначен ректором созданного института. С этого времени в полной мере раскрылись выдающиеся организаторские и педагогические способности Евгения Викторовича. В МИЭМ, благодаря его кипучей и самоотверженной деятельности, были организованы новые кафедры и научные лаборатории, сформирован дружный коллектив единомышленников. В начале 70-ых годов прошлого столетия в МИЭМ по инициативе академика АН СССР С.Н. Вернова и профессора Е.В. Армевского впервые в нашей стране была организована подготовка специалистов в области космического и радиационного материаловедения, созданы лаборатории электронной микроскопии, рентгеновского анализа.

Евгений Викторович Армений уделял большое внимание учебной и научной деятельности. В 1968 году выходит в свет учебное пособие Е. В. Армевского и Г.Б. Фалка

«Электрические микромашины». В 1969 году Евгений Викторович защищает докторскую диссертацию на тему «Системы управления ускорителями заряженных частиц с энергией до 100 Мэв», в этом же году ему присвоено ученое звание «профессор» по кафедре «Электрические машины и электромеханические устройства автоматики». Профессор Евгений Викторович Арменский хорошо известен своими монографиями и учебниками, изданными на русском, английском, французском, испанском языках. Эти книги стали классическими настольными пособиями для специалистов и студентов практически во всех научных и учебных центрах.

В 1973 году Евгению Викторовичу присвоено почетное звание «Заслуженный деятель науки и техники РСФСР». В 1984 году ему вместе с коллегами была присуждена Государственная премия СССР «За создание и внедрение в промышленность сверхвысоковакуумных магниторазрядных насосов и высоковакуумных средств технологического и научного оборудования электронной техники».

Евгений Викторович всегда уделял большое внимание подготовке молодых ученых. Под его руководством выполнены и защищены 10 докторских и свыше 30 кандидатских диссертаций. На протяжении 40 лет Евгений Викторович был членом редколлегии журналов «Измерительная техника», «Физика и химия обработки материалов» и «Датчики и системы».

В своей педагогической деятельности Евгений Викторович Арменский всегда призывал студентов, аспирантов стремиться получить дополнительные знания в области прикладной и фундаментальной математики, а также использовать современные информационные технологии в научной работе.

Секция «Математика и компьютерное моделирование»

ДВУХПАЛУБНАЯ СТРУКТУРА ПОГРАНИЧНОГО СЛОЯ В ТРЕХМЕРНОЙ ЗАДАЧЕ ОБТЕКАНИЯ МАЛОЙ НЕРОВНОСТИ НА ПОВЕРХНОСТИ ПЛАСТИНЫ

Буров Н.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В ходе работы построено асимптотическое решение и получены результаты численного моделирования для задачи обтекания пластины с малой локализованной неровностью несжимаемой вязкой жидкостью с большим числом Рейнольдса на в трехмерном случае.

Введение

Для задач обтекания жидкостями и газами поверхностей с малыми неровностями при больших значениях числа Рейнольдса Re широко известны асимптотические решения с двух- и трехпалубными структурами пограничного слоя [1–5].

Течение вязких жидкостей и газов моделируется при помощи уравнений Навье–Стокса, прямое численное моделирование которых в рамках решения данной задачи является ресурсоемким [6, 7] из-за наличия пространственной разномасштабности (обусловленной геометрией обтекаемой поверхности). Двух- и трехпалубные структуры пограничных слоев позволяют избавиться от этой разномасштабности, и сделать численное моделирование гораздо более эффективным (не требующим применения суперкомпьютеров).

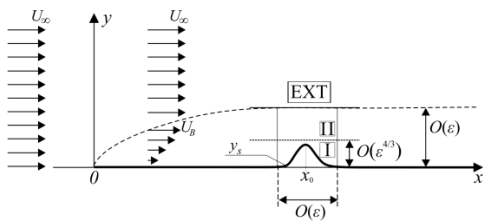


Рис. 1. Схематичное изображение двухпалубной структуры (в сечении): I – тонкий погранслой, II – погранслой Прандтля, EXT – область внешнего (потенциального) потока

Математическая модель рассматриваемой задачи основывается на строгом асимптотическом анализе системы уравнений Навье–Стокса и неразрывности:

$$\langle \mathbf{U}, \nabla \rangle \mathbf{U} = -\nabla p + \varepsilon^2 \Delta \mathbf{U}, \quad \langle \nabla, \mathbf{U} \rangle = 0$$

с граничными условиями прилипания к обтекаемой поверхности y_s , $\mathbf{U}|_{y=y_s} = 0$ и согласованием с набегающим плоскопараллельным потоком $\mathbf{u}_\infty = (1, 0, 0)$ вдали от нее. Здесь $\mathbf{U} = (u, v, w)$ — вектор скорости, p — давление.

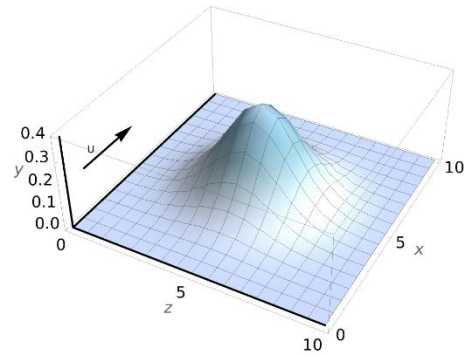


Рис. 2. Малая неровность

Поверхность пластины y_s представлена на Рис. 2 и имеет вид:

$$y_s = \varepsilon^{4/3} \mu((x-x_0)/\varepsilon, (z-z_0)/\varepsilon),$$

$\varepsilon = Re^{-1/2}$ — малый параметр, а μ — некоторая гладкая локализованная в точке (x_0, z_0) функция, убывающая при стремлении аргументов к $\pm\infty$.

Отметим, что решение такой задачи в случае отсутствия неровности выражается через функцию Блазиуса $f(\gamma)$ [8].

Асимптотический анализ

Рассматриваемая задача исследовалась с помощью многомасштабного асимптотического анализа, основанного на комбинации метода погранслоевых разложений и метода построения локализованных решений [2, 3], и полученные результаты можно сформулировать в виде следующей теоремы.

Теорема. Пусть $x_0, z_0 \geq \delta > 0$. Тогда:

$$u = 1 + u_0^{\text{II}}(x, \tau) + \varepsilon^{1/3} (u_1^{\text{I}}(\xi_1, \xi_2, \theta) + u_1^{\text{II}}(\xi_1, \xi_2, \tau)) + O(\varepsilon^{2/3}),$$

$$v = \varepsilon^{2/3} (v_2^{\text{I}}(\xi_1, \xi_2, \theta) + v_2^{\text{II}}(\xi_1, \xi_2, \tau)) + O(\varepsilon),$$

$$w = \varepsilon^{1/3} w_1^{\text{I}}(\xi_1, \xi_2, \theta) + O(\varepsilon^{2/3}),$$

$$p = p_0 + \varepsilon^{2/3} p_2^{\text{II}}(\xi_1, \xi_2, \tau) + O(\varepsilon),$$

$$\text{где } \xi_1 = \frac{x-x_0}{\varepsilon}, \xi_2 = \frac{z-z_0}{\varepsilon}, \tau = \frac{y_w}{\varepsilon}, \theta = \frac{y_w}{\varepsilon^{4/3}},$$

$y_w = y - y_s$ — переменная, которая выравнивает границу, τ и θ — погранслоевые переменные для II и I

палуб соответственно (EXT, I, II - обозначения палуб потока согласно Рис. 1).

Функция $u_0^{\text{II}} = u^* - 1$, $u_1^{\text{II}} = \mu \frac{\partial u_0^{\text{II}}}{\partial \tau} \Big|_{x=x_0}$, где

$u^* = f' \left(\frac{\tau}{\sqrt{x}} \right)$, $f(\gamma)$ - функция Блазиуса, которая

является решением следующей краевой задачи:

$$f''' = -\frac{1}{2} f f'', \quad f(0) = f'(0) = 0, \quad f'(\infty) = 1$$

Функции $u_1^{\text{I}}, v_2^{\text{I}}, w_1^{\text{I}}$ определяются из соотношений

$$u^\dagger = u_1^{\text{I}} + u_1^{\text{II}} \Big|_{\tau=0} + \theta \frac{\partial u_0^{\text{II}}}{\partial \tau} \Big|_{\tau=0}, \quad v^\dagger = v_2^{\text{I}} + v_2^{\text{II}} \Big|_{\tau=0}, \quad w^\dagger = w_1^{\text{I}},$$

где функции $u^\dagger, v^\dagger, w^\dagger$ являются решением краевой задачи для системы уравнений Прандтля с самоиндуцированным давлением:

$$\begin{cases} u^\dagger \left(\frac{\partial u^\dagger}{\partial \xi_1} - \frac{\partial \mu}{\partial \xi_1} \frac{\partial u^\dagger}{\partial \theta} \right) + w^\dagger \left(\frac{\partial u^\dagger}{\partial \xi_2} - \frac{\partial \mu}{\partial \xi_2} \frac{\partial u^\dagger}{\partial \theta} \right) + \\ v^\dagger \frac{\partial u^\dagger}{\partial \theta} + \frac{\partial p_2^{\text{II}}}{\partial \xi_1} \Big|_{\tau=0} - \frac{\partial^2 u^\dagger}{\partial \theta^2} = 0, \\ v^\dagger \frac{\partial w^\dagger}{\partial \theta} + w^\dagger \left(\frac{\partial w^\dagger}{\partial \xi_2} - \frac{\partial \mu}{\partial \xi_2} \frac{\partial w^\dagger}{\partial \theta} \right) + u^\dagger \left(\frac{\partial w^\dagger}{\partial \xi_1} - \frac{\partial \mu}{\partial \xi_1} \frac{\partial w^\dagger}{\partial \theta} \right) - \\ - \frac{\partial^2 w^\dagger}{\partial \theta^2} = 0, \\ \frac{\partial u^\dagger}{\partial \xi_1} - \frac{\partial \mu}{\partial \xi_1} \frac{\partial u^\dagger}{\partial \theta} + \frac{\partial w^\dagger}{\partial \xi_2} - \frac{\partial \mu}{\partial \xi_2} \frac{\partial w^\dagger}{\partial \theta} + \frac{\partial v^\dagger}{\partial \theta} = 0, \end{cases} \quad (1)$$

с граничными условиями:

$$\begin{aligned} u^\dagger \Big|_{\theta=0} = 0, \quad v^\dagger \Big|_{\theta=0} = 0, \quad w^\dagger \Big|_{\theta=0} = 0, \\ \frac{\partial u^\dagger}{\partial \theta} \Big|_{\theta \rightarrow \infty} = \frac{\partial u_0^{\text{II}}}{\partial \tau} \Big|_{\tau=0}, \quad u^\dagger \Big|_{\xi_{1,2} \rightarrow \pm \infty} = \theta \frac{\partial u_0^{\text{II}}}{\partial \tau} \Big|_{\tau=0}, \\ w^\dagger \Big|_{\theta \rightarrow \infty} = 0, \quad v^\dagger \Big|_{\xi_{1,2} \rightarrow \pm \infty} = 0, \quad w^\dagger \Big|_{\xi_{1,2} \rightarrow \pm \infty} = 0. \end{aligned} \quad (2)$$

Функция v_2^{II} является решением краевой задачи для уравнения типа Рэлея, возникающего в системе

$$\Delta_{\xi_1, \xi_2, \tau} = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2} + \frac{\partial^2}{\partial \tau^2} :$$

$$\begin{aligned} u^* \Big|_{x=x_0} \Delta_{\xi_1, \xi_2, \tau} v_2^{\text{II}} - v_2^{\text{II}} \frac{\partial^2 u^*}{\partial \tau^2} \Big|_{x=x_0} = 0, \\ v_2^{\text{II}} \Big|_{\xi_{1,2} \rightarrow \pm \infty} = 0, \quad v_2^{\text{II}} \Big|_{\tau \rightarrow \infty} = 0, \quad v_2^{\text{II}} \Big|_{\tau=0} = v^\dagger \Big|_{\theta \rightarrow \infty} \end{aligned}$$

Давление p_2^{II} определяется равенством:

$$p_2^{\text{II}} = \int_{-\infty}^{\xi_2} \frac{\partial w_2^{\text{II}}}{\partial \xi_1} d\xi_2',$$

где функция w_2^{II} является решением краевой задачи

$$\Delta_{\xi_1, \xi_2} = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2} :$$

$$\begin{aligned} u^* \Big|_{x=x_0} \Delta_{\xi_1, \xi_2} w_2^{\text{II}} = \frac{\partial v_2^{\text{II}}}{\partial \xi_2} \frac{\partial u^*}{\partial \tau} \Big|_{x=x_0} - \frac{\partial^2 v_2^{\text{II}}}{\partial \xi_2 \partial \tau} u^* \Big|_{x=x_0}, \\ w_2^{\text{II}} \Big|_{\xi_{1,2} \rightarrow \pm \infty} = 0. \end{aligned}$$

Выражение для самоиндуцированного давления можно записать в виде:

$$\frac{\partial p_2^{\text{II}}}{\partial \xi_1} \Big|_{\tau=0} = -v^\dagger \Big|_{\theta \rightarrow \infty} \frac{\partial u_0^{\text{II}}}{\partial \tau} \Big|_{\tau=0}.$$

Это равенство позволяет решать систему уравнений Прандтля с самоиндуцированным давлением (1) независимо от остальных уравнений - нужно знать лишь значение $f''(0)$, которое известно и приближенно равно 0.33.

Численное моделирование течения вблизи трехмерной неровности

Объектом численного моделирования является система уравнений (1) с граничными условиями (2).

Для численного моделирования был использован метод конечных разностей. Решения задачи (1) стационарные, для их поиска используется метод установления. Устойчивость схемы была улучшена применением техники выбора разности по потоку или против потока в зависимости от знака коэффициента перед производной.

Программная реализация алгоритмов моделирования выполнена на языке C++ с применением библиотек VTK (для визуализации) и OpenMP (для распараллеливания явной разностной схемы).

Для моделирования течения была выбрана неровность с высотой $\theta=1$ с координатами $x_0=1, z_0=1$.

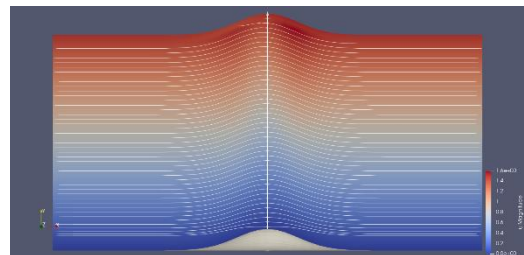
Прочие параметры:

1. $N=51, M=51, K=51$ – количество узлов сетки в направлениях Ox, Oy, Oz .
2. Шаг виртуального времени $1/4000$.
3. Область вычислений: $\xi_{1,2} \in [-10, 10], \theta \in [0, 10]$

Согласно работе [9] ошибка от ограничения области вычислений сравнима с ошибкой от аппроксимации.

4. Точность, до которой выполняются вычисления: 10^{-5} (т. е. условие что полученное решение является стационарным – сеточная норма разницы между двумя итерациями меньше этой точности).

Неожиданным при моделировании является практически отсутствие обтекания потоком неровности слева и справа. Вытеснение потока неровностью происходит преимущественно вверх (см. Рис. 3. Визуализация результатов численного моделирования).



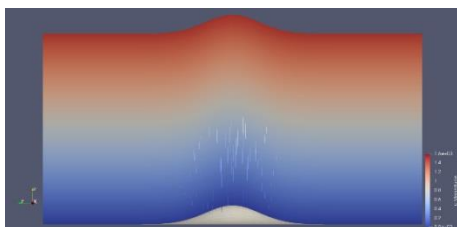


Рис. 3. Визуализация результатов численного моделирования

Заключение

В ходе работы были получены асимптотики для решения стационарной задачи обтекания набегающим потоком плоскости с малой локализованной неровностью на ее поверхности в трехмерном случае.

Для задачи обтекания плоскости с малой локализованной неровностью на ее поверхности в трехмерном случае проведено численное моделирование при помощи методов разностных схем на языке C++ с применением библиотек OpenMP и VTK. Неожиданным результатом в ходе работы стало отсутствие обтекания слева и справа потоком малой локализованной неровности в трехмерном случае.

Список литературы

1. Danilov V. G. Double-deck structure of the boundary layer in problems of flow around localized perturbations on a plate / V. G. Danilov, R. K. Gaydukov // *Mathematical Notes*. – 2015. – № 98. – С. 561–571.
2. Gaydukov R. K. Double-deck structure of the boundary layer in the problem of a compressible flow along a plate with small irregularities on the surface. / R. K. Gaydukov // *European Journal of Mechanics - B/Fluids*. – 2017. – № 66. – С. 102–108.
3. Gaydukov R. K. Double-deck structure in the fluid flow problem over plate with small irregularities of time-dependent shape. / R. K. Gaydukov // *European Journal of Mechanics-B/Fluids*. – 2021. – № 89. – С. 401–410.
4. Smith F. T. Laminar flow over a small hump on a flat plate / F. T. Smith // *Journal of Fluid Mechanics*. – 1973. – № 57. – С. 803–824.
5. Yapalparvi R. Double-deck structure revisited / R. Yapalparvi // *European Journal of Mechanics - B/Fluids*. – 2012. – № 31. – С. 53–70.
6. Direct numerical simulation of a compressible boundary-layer flow past an isolated three-dimensional hump in a high-speed subsonic regime / D. De Grazia, D. Moxey, S. J. Sherwin [и др.] // *Phys. Rev. Fluids*. – 2018. – № 3. – С. 024101.
7. Triple-deck and direct numerical simulation analyses of high-speed subsonic flows past a roughness element. / M. Gianmarco, M. Kravtsova, A. Ruban, S. Sherwin // *Journal of Fluid Mechanics*. – 2015. – № 774. – С. 311–323.
8. Шлихтинг Г. Теория пограничного слоя / Г. Шлихтинг. – Москва : Наука, 1974. – 712 с.
9. Gaydukov R. K. Numerical algorithm for solving prandtl equations with induced pressure in the periodic case. / R. K. Gaydukov // *Сибирский журнал вычислительной математики*. – 2022. – Т. 15, № 2. – С. 97–108.

МИНИМАЛЬНОЕ ПОКРЫВАЮЩЕЕ МНОЖЕСТВО КАК ИНСТРУМЕНТ ОПТИМАЛЬНОГО КОЛЛЕКТИВНОГО ВЫБОРА

Юдина А.В.

Национальный исследовательский университет
«Высшая школа экономики»
Высшая школа бизнеса

Аннотация

В работе рассматриваются способы выбора наилучших альтернатив на основании результатов их попарного сравнения. Подобный выбор является проблемой в ситуации, когда у любого варианта выбора есть более предпочтительный вариант. В литературе предложено много концепций решения поставленной задачи (так называемых турнирных решений). При этом особый интерес представляют обобщения турнирных решений на случай неполных сравнений, так как реальные данные обычно неполны.

Настоящее исследование посвящено обобщениям минимального покрывающего множества. Актуальность работы обусловлена недостаточностью имеющегося описания аксиоматических свойств рассматриваемых версий этого решения и их теоретико-множественных соотношений с другими турнирными решениями, что препятствует их использованию в прикладных задачах.

Работа имеет как теоретическое, так и практическое значение. В рамках проведенного исследования были описаны свойства версий минимального покрывающего множества и их теоретико-множественные соотношения с другими турнирными решениями. Также в работе представлены результаты решения прикладной задачи - ранжирования научных журналов. В ходе исследования была разработана программа на языке Python [1], автоматизирующая ряд рутинных операций, таких как вычисление турнирного решения, проведение рангового корреляционного анализа и составление аналитической отчетности.

Введение

В экономической теории одним из центральных является понятие полезности экономического блага. Рассматриваемый в данной работе способ ранжирования благ сортировкой, основанной на выборе с помощью турнирного решения, является функцией общественного благосостояния. Концепция функции общественного благосостояния, предложенная Кеннетом Эрроу [2], была логическим развитием идей так называемой «новой» экономики общественного благосостояния (new welfare economics). В свою очередь, «новая» экономическая теория была критикой и продолжением «старой» экономики общественного благосостояния, созданной трудами Рикардо, Мальтуса и Дж.Ст. Милля [3, 4, 5]. Основной целью этой «старой» теории было объективное определение и измерение благосостояния общества (или любого иного субъекта коллективного выбора). В качестве

такой меры было предложено суммарное богатство, то есть сумма индивидуальных полезностей или выгод, измеряемых в денежном эквиваленте. Данный подход получил название утилитаризма. С точки зрения утилитариста индивидуальные полезности являются не только количественно измеримыми, но и сравнимыми, а перераспределение экономических благ трансферами не меняет общественного благосостояния.

«Новые» экономисты (Самуэльсон, Литтл и другие [6, 7]) поставили под сомнение утилитаристский тезис о сравнимости индивидуальных полезностей. Даже если индивидуальные блага могут быть выражены в денежном исчислении, мы не можем предполагать индивидуальную полезность доллара в карманах богача и бедняка одинаковой. Кардинализм утилитаристов предлагалось заменить ординальной интерпретацией функций полезности. Сравнить полезности разных индивидов запрещалось. По сути, вместо функций полезности было предложено рассматривать бинарные отношения и строить предпочтения коллектива на основании профиля этих отношений с помощью того или иного правила агрегирования, именуемого функцией общественного благосостояния. Очевидно, что эти правила должны были использовать не кардинальные операции, как взвешенная сумма или взвешенное произведение полезностей, а ординальные, например, сравнение по правилу большинства.

Однако затем Кеннет Эрроу [2] показал, что коллективные предпочтения нельзя представить даже ординальной функцией полезности, поскольку репрезентирующее их отношение не является в общем случае транзитивным. Нетранзитивность же отношения предпочтения приводит к тому, что максимальных элементов этого отношения может просто не существовать. Данный факт исключает наивное определение оптимизации как максимизации, то есть как выбор максимального элемента отношения предпочтения.

Возникает проблема оптимального выбора в ситуации противоречивых предпочтений, то есть проблема выбора в некотором (не наивном) смысле лучших альтернатив на основании результатов их попарного сравнения, репрезентированных нетранзитивным отношением. Эту проблему можно представить как задачу выбора победителя или победителей в спортивном турнире, и потому концепции её решения были названы *турнирными решениями*.

В литературе было предложено довольно много различных турнирных решений. Свойства этих решений, в целом, хорошо исследованы в случае, когда отношение строгого предпочтения является полным, то есть когда нет различных альтернатив, равноценных для выбирающего. Особенная важность придаётся таким свойствам, как монотонность и устойчивость выбора. Данными свойствами обладают далеко не все известные турнирные решения, соответственно монотонные и устойчивые решения имеют особую ценность. Именно благодаря монотонности и устойчивости, одним из кандидатов на роль наилучшего решения является так называемое *минимальное*

покрывающее множество (minimal covering set), предложенное Дутта [8]. Концепция этого решения опирается на усиление отношения строгого предпочтения, называемое отношением покрытия.

Как было отмечено ранее, в прикладных задачах часты ситуации, когда в актуальном меню выбора есть пары равноценных альтернатив.¹ Эти ситуации соответствуют турнирам, где разрешены ничьи. Соответственно, есть практическая необходимость обобщить турнирные решения, предложенные для полных отношений, на случай выбора победителя в турнирах с ничьими.

Работа посвящена рассмотрению минимального по МакКельви множества, минимального по Дугану множества и консервативного обобщения минимального покрывающего множества, предложенных в качестве естественных обобщений минимального покрывающего множества, которые затем применяются в задаче ранжирования научных журналов.

Минимальное покрывающее множество и его обобщения на случай слабых турниров

Определение минимального покрывающего множества основано на отношении покрытия, для определения которого нам потребуется ввести понятия верхнего и нижнего срезов и горизонта альтернативы.

Верхний срез $D(x)$ альтернативы x есть множество всех альтернатив, которые строго предпочтительнее x , $D(x) = \{y \in X | y \succ x\}$. Нижний срез $L(x)$ альтернативы x есть множество всех альтернатив, которым строго предпочитается x , $L(x) = \{y \in X | x \succ y\}$. Горизонт $H(x)$ альтернативы x есть множество всех альтернатив равноценных x , $H(x) = \{y \in X \setminus \{x\} | x \sim y\}$.

Говорят, что альтернатива x покрывает альтернативу y (обозначим это отношение как $x \text{Cu} y$), если x не только лучше, чем y , но и лучше любой альтернативы, которая хуже, чем y , $x \text{Cu} y \Leftrightarrow (x \succ y) \wedge (L(y) \subset L(x))$. Альтернатива называется непокрытой тогда и только тогда, когда в предъявлении нет ни одной покрывающей её альтернативы. Непокрытое множество UC есть множество всех непокрытых альтернатив предъявления, $UC = \{x \in X | \forall y \in X, \neg y \text{Cu} x\}$ [9].

Если рассматриваемый турнир является полным, покрывающим множеством $Y \subseteq X$ называют множество, удовлетворяющее следующим условиям:

- 1) $UC(Y) = Y$
- 2) $\forall x \in X \setminus Y, x \notin UC(Y \cup \{x\})$

В связи с тем, что субъект выбора обычно предпочитает сузить множество выбираемых альтернатив, оптимальным решением будет являться минимальное по включению покрывающее множество. Известно, что такое множество будет единственным. Получившее турнирное решение было названо минимальным покрывающим множеством MC [9].

В слабом турнире существует 15 различных версий отношения покрытия, но нам понадобятся только две из них, а именно: версия МакКельви и версия

¹ Несравнимые альтернативы без потери общности можно считать равноценными.

Дуггана. Альтернатива x покрывает альтернативу y по МакКельви (обозначим данное отношение как $x C_{McK} y$), если $L(y) \subset L(x) \wedge D(x) \subset D(y)$. Альтернатива x покрывает альтернативу y по Дуггану, если x лучше всех альтернатив, которые хуже или равноценны y , $x C_D y \Leftrightarrow H(y) \cup L(y) \subset L(x)$. По аналогии с определением непокрытого множества, непокрытое по версии V множество UC_V (где $V \in \{McK, D\}$) представляет из себя объединение всех непокрытых по версии V альтернатив предъявления, $UC_V = \{x \in X | \forall y \in X, \neg y C_V x\}$ [9].

Минимальное покрывающее множество можно естественным образом обобщить путем замены непокрытого множества в исходном определении на некоторую его версию. Таким образом мы получаем минимальное покрывающее по МакКельви множество MC_{McK} и минимальное покрывающее по Дуггану множество MC_D .

Кроме того, минимальное покрывающее множество можно обобщить с помощью введенного Ф. Брандтом понятия консервативного обобщения [10]. Для его получения нужно составить множество ориентаций $[W]$ слабого турнира $W(X, \succeq)$, то есть множество всех турниров $T(X, >)$, которые можно получить произвольным дополнением слабого турнира $W(X, \succeq)$ до полного. Консервативным обобщением минимального покрывающего множества $[MC]$ называется множество альтернатив, выбираемых минимальным покрывающим множеством из всех турниров, входящих в множество ориентаций $[W]$ слабого турнира W ,

$$[MC](W) = \bigcup_{T \in [W]} MC(T).$$

Свойства версий минимального покрывающего множества и их связь с другими турнирными решениями

В рамках подхода, предложенного Кеннетом Эрроу, оптимальность выбора оценивается по наличию у принятого правила выбора ряда естественных свойств, характеризующих его как рациональную процедуру, в частности, по наличию свойств монотонности, устойчивости и вычислительной простоты.

Известно, что устойчивой и монотонной является версия минимального покрывающего множества MC_{McK} , соответствующая версии отношения покрытия, предложенной МакКельви [11]. Для версии Дуггана MC_D была доказана лишь устойчивость. Однако, используя соответствующее рассуждение для MC_{McK} , нетрудно показать, что MC_D также удовлетворяет свойству монотонности.

В [12] предложен полиномиальный алгоритм для вычисления минимального покрывающего по МакКельви множества. В ходе данного исследования показано, что этот алгоритм (с точностью до перепределения отношения покрытия) можно использовать и для нахождения MC_D .

В силу того, что консервативное обобщение наследует свойства исходного турнирного решения [10], был сделан вывод, что консервативное обобщение мини-

мального покрывающего множества $[MC]$ удовлетворяет свойствам монотонности и устойчивости.

Несмотря на хорошие теоретические свойства, возможности использования консервативного обобщения минимального покрывающего множества сильно ограничены, так как неизвестен полиномиальный алгоритм вычисления $[MC]$.

Помимо свойств трёх версий минимального покрывающего множества были исследованы их теоретико-множественные соотношения с различными обобщениями других турнирных решений, а именно: с обобщениями двухпартийного множества BP (E и $[BP]$), максимального цикла TC (STC , WTC и UT), непокрытого множества UC (UC_{IM} , UC_M , UC_{IF} , UC_F , UC_{McK} и UC_D), объединения минимальных доминирующих множеств D (D , SP и $[D]$), объединения минимальных внешне устойчивых множеств ES (WS , ES , RES и $[ES]$) и незахваченного множества UCp (UCp). Определения перечисленных решений можно найти в [9, 13]. Установленные факты перечислены ниже в виде теоремы.

Теорема 1.

1) MC_{McK} всегда является подмножеством MC_D , UC_{McK} , UC_D , $[D]$, $[ES]$, UCp , WTC . MC_{McK} всегда является надмножеством E . MC_{McK} не включено в $[BP]$, $[MC]$, UC_M , UC_{IF} , UC_F , D , SP , WS , ES , RES , STC , UT и наоборот, но их пересечение всегда непусто. MC_{McK} не включено в UC_{IM} и наоборот, причем их пересечение может быть пустым.

2) MC_D всегда является подмножеством UC_D , $[D]$, $[ES]$, WTC . MC_D всегда является надмножеством E , $[BP]$, MC_{McK} , $[MC]$. MC_D не включено в UC_{IM} , UC_M , UC_{IF} , UC_F , UC_{McK} , D , SP , WS , ES , RES , UCp , STC , UT и наоборот, но их пересечение всегда непусто.

3) $[MC]$ всегда является подмножеством MC_D , UC_D , $[D]$, $[ES]$, WTC . $[MC]$ всегда является надмножеством $[BP]$. $[MC]$ не включено в MC_{McK} , UC_{IM} , UC_M , UC_{IF} , UC_F , UC_{McK} , D , SP , WS , ES , RES , UCp , STC , UT и наоборот, но их пересечение всегда непусто.

Доказательство Теоремы 1 дано в [14]. Следует отметить, что для пары $[MC]$ и E результат сравнения пока не известен и является открытой проблемой.

Эмпирическое сравнение методов ранжирования

Для оценки эффективности рассмотренных решений как инструментов выбора проведено сравнение результатов их применения в прикладной задаче с результатами, которые дают другие турнирные решения. Сравнение основано на ранговом корреляционном анализе ранжирований научных журналов по менеджменту. В качестве меры корреляции использовался коэффициент Кендалла τ_b . Методология сравнения и интерпретации результатов заимствована из [15].

Исходными данными были пять ранжирований 177 научных журналов по менеджменту, опубликованные в [1]. Рейтинги были построены по значениям пяти наиболее популярных библиометрических показателей (двухлетнему импакт-фактору, пятилетнему импакт-фактору, индексу влияния статьи, SNIP

и SJR), рассчитанным для 2022 года. В рассматриваемой совокупности журналов имеется 1264 равноценные пары.

Пять исходных ранжирований агрегировались в (нетранзитивное) мажоритарное отношение с помощью правила большинства. По мажоритарному отношению сортировкой, основанной на одном из турнирных решений, строилось агрегированное ранжирование. Ранжирования, исходные и агрегированные, сравнивались попарно и упорядочивались следующим образом. Вслед за авторами [15] предполагалось, что ранжирование I репрезентирует множество агрегируемых ранжирований лучше ранжирования II , если ранжирование I коррелирует с большинством из агрегируемых ранжирований сильнее, чем ранжирование II . Ранжирование I коррелирует с ранжированием III сильнее согласно, например, коэффициенту τ_b , чем ранжирование II , если значение τ_b для пары $\{I, III\}$ больше значения τ_b для пары $\{II, III\}$. С помощью попарного сравнения было определено мажоритарное отношение на множестве ранжирований. Таблица со значениями коэффициента τ_b , а также полученная на ее основе матрица строгого мажоритарного отношения приведены в [1].

Для ранжирования имеющихся ранжирований к полученному мажоритарному отношению была применена вторая версия правила Коупланда, согласно которой альтернативы были расставлены в порядке убывания значений оценок Коупланда s_2 , где $s_2(x) = |L(x)|$. Ранжирование ранжирований, полученное на основе этих оценок, приведено в Таблице 1.

Таблица 1. Ранжирование ранжирований

Ранг	Метод
1	Сортировка с помощью MC_D
2	Правило Коупланда (2 версия)
3	Сортировка с помощью ES
4	Правило Коупланда (3 версия) / Сортировка с помощью UC_M / Сортировка с помощью MC_{MCK}
5	
6	
7	Сортировка с помощью E
8	5-летний импакт-фактор
9	SNIP
10	Импакт-фактор (2-летний)
11	SJR
12	Индекс влияния статьи

Примечание. В таблице 1 ранжирования ранжируются по тому, насколько хорошо они могут репрезентировать исходную совокупность из пяти агрегируемых ранжирований. Методы расположены в порядке ухудшения качества. Источник: составлено автором на основе результатов работы программы [1].

Заключение

В ходе проведенного исследования были рассмотрены важные свойства трёх версий минимального покрывающего множества, такие как монотонность, устойчивость и вычислительная простота. Было установлено, что все три версии сохраняют эти

свойства исходного турнирного решения в случае, когда есть пары равноценных альтернатив.

Также была исследована связь трех версий минимального покрывающего множества с многими другими известными обобщениями турнирных решений на случай слабых турниров.

Кроме того, была решена прикладная задача, связанная с управлением в академической сфере, – агрегирование противоречивых рейтингов научных журналов (по менеджменту), основанных на различных библиометрических показателях. Задача решалась на актуальных данных, опубликованных в 2022 году. Проведенный корреляционный анализ показал, что обсуждаемые обобщения минимального покрывающего множества представляют исходную совокупность агрегируемых рейтингов лучше, чем каждый из них, а в случае минимального покрывающего по Дугану множества, еще и лучше, чем агрегированные ранжирования, построенные с помощью иных используемых в работе процедур выбора.

Для выполнения исследования была написана программа на языке программирования Python [1]. Разработанное программное решение позволило оптимизировать ряд рутинных операций, таких как предобработка данных, расчет турнирных решений, построение с их помощью ранжирований, проведение корреляционного анализа, построение и вывод аналитической отчетности.

Благодарности

Автор выражает благодарность своему научному руководителю Субочеву Андрею Николаевичу, кандидату физико-математических наук, доценту департамента математики факультета экономических наук НИУ ВШЭ, за постановку задачи, важные советы при проведении исследования и рекомендации по оформлению данного текста.

В данной работе представлены результаты, полученные в рамках реализации проекта «Теоретические и прикладные задачи анализа данных и принятия решений в социально-экономической и финансовой сферах», выполненного в рамках Программы фундаментальных исследований НИУ ВШЭ в 2022 году. Исследование проводилось при финансовой поддержке Международного центра анализа и выбора решений НИУ ВШЭ.

Список литературы

1. Репозиторий проекта на GitHub [Электронный ресурс]. URL: <https://github.com/AngelinaYudina/Journals> (дата обращения: 27.04.2023).
2. Arrow, K.J. Social choice and individual values. / K.J. Arrow // Yale University Press. – 1951. – p. 99.
3. Malthus, T.R. Principles of Political Economy. / T.R. Malthus // Cambridge University Press. – 1989. – p. 1147.
4. Mill, J.S. Utilitarianism. / J.S. Mill, R. Crisp (ed.) // Oxford University Press. – 1998. – p. 168.
5. Ricardo, D. On The Principles of Political Economy and Taxation. / D. Ricardo // John Murray. – 1817. – p. 589.
6. Little, I.M.D. A critique of welfare economics. / I.M.D. Little // Clarendon, Oxford. – 1957. – p. 324.

7. Samuelson, P.A. Bergsonian welfare economics. / P.A. Samuelson Текст: непосредственный // Economic welfare and the economics of Soviet Socialism. / P.A. Samuelson, S. Rosefielde (ed) – Cambridge University Press. – 1981. – p. 223-266.
8. Dutta, B. Covering sets and a new Condorcet choice correspondence. / B. Dutta // Journal of Economic Theory. – 1988. – Vol. 44. – p. 63–80.
9. Brandt, F. Tournament solutions. / F. Brandt, M. Brill, P. Harrenstein // Cambridge University Press. – 2016. – p. 57-84.
10. Brandt, F. Extending tournament solutions. / F. Brandt, M. Brill, P. Harrenstein // Social Choice and Welfare. – 2018. – Vol. 51(2). – p.193–222.
11. Peris, J.E. Condorcet choice correspondences for weak tournaments. / J.E. Peris, B. Subiza // Social Choice and Welfare. – 1999. – Vol. 16(2). – p. 217–231.
12. Brandt, F. Computing the minimal covering set. / F. Brandt, F. Fischer // Mathematical Social Sciences. – 2008. – Vol. 56(2). – p. 254–268.
13. Subochev, A. Dominating, weakly stable, and uncovered sets: Properties and generalizations. / A. Subochev // Automation and Remote Control. – 2010. – Vol. 71(1). – p. 116-127.
14. 2023. Subochev, A., Yudina, A. Old and new tournament solutions based on choosing the best subsets of a menu (forthcoming).
15. Subochev, A. Ranking journals using social choice theory methods: A novel approach in bibliometrics. / F. Aleskerov, V. Pisyakov, A. Subochev // Journal of Informetrics. – 2018. – No. 12(2). – p. 416–429.

ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ КРИТИЧЕСКОГО ПОВЕДЕНИЯ В МОДЕЛИ БЛЮМЕ-КАПЕЛЯ МЕТОДОМ ПРЯМОЙ ОЦЕНКИ ЭНТРОПИИ

Мозоленко В.К.

*Национальный исследовательский университет
«Высшая школа экономики»,
аспирантская школа по физике*

Аннотация

В работе предложена модификация метода Розе-Махты прямого расчета энтропии. Оригинальный метод позволял определить энтропию на нижней части спектра энергий исследуемой системы. Наша модификация дает возможность рассчитать энтропию на всем спектре энергий. Это дает возможность получить полную информацию о статистической сумме и ее производных по температуре. Которые, в свою очередь, позволяют вычислить температурную зависимость внутренней энергии и удельной теплоемкости. Модифицированный метод применен к исследованию фазовых переходов модели Блюме-Капеля.

Введение

Фазовые переходы характеризуются особенностью в первой или второй производной свободной энергии. Такие переходы называются фазовыми переходами первого и второго рода, соответственно.

Численное исследование фазовых переходов затруднено в обоих случаях, по причинам, связанным с физикой переходов [1].

Для фазового перехода первого рода затруднением становится существование границы разделения фаз, вероятность перехода через которую для локальных алгоритмов Монте-Карло, наиболее часто используемых, мала. Энергетический барьер между фазами растет линейно с размером системы L , а время туннелирования между фазами растет как $\exp(L)$, что приводит к экспоненциальному росту необходимого числа шагов алгоритма с ростом размера системы.

Для фазового перехода второго рода затруднением для применения локальных алгоритмов Монте-Карло становится явление критического замедления, связанного с ростом корреляционной длины параметра порядка по мере приближения к фазовому переходу. Что приводит к степенному росту времени релаксации с размером системы.

Развитие и исследование методов, которые позволяют преодолеть эти трудности, являются актуальными задачами. Данная работа посвящена численному исследованию с применением нового метода.

Постановка задачи

В настоящей работе предлагается модификация метода, предложенного Натаном Розе и Джоном Махтой в работе [2] для исследования окрестности фазового перехода первого рода, и используем его для исследования модели Блюме-Капеля [3,4]. Предложенный в работе [2] метод отсечения образцов с энергиями выше некоторого потолка энергии позволял получить оценку энтропии, однако, только энтропии, соответствующей левой ветви энергетического спектра, что не позволяет сделать оценку термодинамических величин, таких как внутренняя энергия и теплоемкость.

Метод модифицирован с целью получения всей функции плотности энергии и использовании этой функции для вычисления термодинамических величин для систем с дискретной энергией. Знание этих величин позволяет определить критические индексы и температуру перехода в термодинамическом пределе бесконечного размера решетки. Для этого используется известный и проверенный прием нахождения асимптотики при знании этих величин при нескольких конечных размерах решетки (finite-size scaling).

Разработанный метод применен к модели Блюме-Капеля [3,4], имеющей сложную фазовую диаграмму с двумя ветвями фазовых переходов (первого и второго рода) и трикритической точкой. Проведено сравнение с результатами применения других методов для исследования этой модели.

Модель Блюме-Капеля представляет собой обобщение модели Изинга со спином 1, определяемая гамильтонианом

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j + D \sum_i \sigma_i^2,$$

где спиновые переменные σ_i принимают три значения -1, 0, +1, а обозначение $\langle i, j \rangle$ задает суммирование по всем соседним спином на двумерной решетке.

Описание алгоритма

Метод Махты и Розе позволяет вычислять непосредственно энтропию системы. На каждом шаге моделирования большого числа R копий системы производится независимый переход каждой копии в новое состояние, но так, чтобы энергия каждой копии системы не превышала потолка энергии. После некоторого большого числа параллельных преобразований системы значение потолка энергии уменьшается, и подсчитывается доля $\epsilon(U)$ отбрасываемых копий, чья энергия оказывается выше нового потолка. По этой доле $\epsilon(U)$ производится оценка энтропии, как функции текущего значения потолка энергии. Реплики с энергиями, меньшими значения нового потолка энергии, случайным образом копируются на освободившиеся места до достижения первоначального числа копий R . Таким образом, сдвигая потолок с одного уровня энергии на другой, получаем значения энтропии, как функции энергии.

Детали алгоритма следующие. Моделируется равновесная физическая система в микроканоническом ансамбле. Рассмотрим $R \gg 1$ реплик физической (например, спиновой) системы, для которой известен либо сам спектр энергий, либо минимальный шаг спектра.

1. Выбирается начальное значение потолка энергии U .

2. Производятся изменения исследуемой системы в конфигурационном пространстве за счет локальных и случайных изменений спинов, что приводит к изменению энергии каждой отдельной копии системы, в течение времени T . При этом допускаются только те принятия изменений спинов, который оставляют копию системы под потолком U .

3. Потолок U сдвигается на следующую по спектру энергию; R' реплик с $E_i \geq U$, оказавшихся выше нового порога, отсеиваются, а их доля от общего числа реплик записывается как функция потолка $\epsilon(U) = R'/R$.

4. Оставшиеся $R - R'$ реплик размножаются случайным образом до R для повторения алгоритма с шага 1.

Этот цикл повторяется до достижения $U = E_{min}$. Зная $\epsilon(U)$ и вырожденность основного состояния, можно восстановить плотность состояний снизу вверх по энергетическому спектру.

Энтропия системы $S(E)$ может быть вычислена как: $S(E) = \ln(\epsilon(E)) + \sum_{E' > E} \ln(1 - \epsilon(E'))$

Однако, таким образом, при применении оригинального метода Розе-Махты, возможно

получить лишь “нижнюю”, низкоэнергетическую часть спектра энергий.

Для того чтобы получить “верхнюю” часть спектра, мы предлагаем “инвертировать” данный алгоритм, уже с энергетическим полом (вместо потолка) и шить результаты по пересекающейся части спектра. Полученная таким образом энтропия определена на всем спектре энергий.

Результаты моделирования

Объектом для исследования точности и применимости метода была выбрана модель Блюме-Капеля [3,4], которая испытывает как фазовый переход первого рода, так и фазовый переход второго рода в зависимости от параметра D , а также особое поведение к трикритической точке (ТСР). Было произведено численное моделирование с целью получения значений энтропии модели Блюме-Капеля при значениях отношения параметров гамильтониана D/J , указанных в таблице 1. При каждом отношении D/J было рассчитано 131 072 копий системы, причем как для моделирования с энергетическим потолком, так и для моделирования с энергетическим полом. Моделирование было проведено для каждого линейного размера решетки от 8 до 64. Полученные значения энтропии были использованы для вычисления удельной теплоемкости, по максимуму которой была проведена оценка критической температуры с использованием метода конечномерного анализа.

В таблице 1 приведены значения оценки критической температуры и сравнение с результатами других авторов, полученными методом MCMC [5], методом Ванга-Ландау [6], методом исследования низко- и высокотемпературных рядов [7] и методом phenomenological finite-size scaling [8]. Эти методы выбраны как популярные и проверенные временем.

Таблица 1. Сравнение критических температур, полученных разными методами для модели Блюме-Капеля. Легенда: FFS - finite size-scaling, MCMC - Monte-Carlo Markov Chain, WL - метод Ванга-Ландау, MRM – модифицированный метод Розе-Махты, TO - род фазового перехода.

D/J	FFS [8]	MCM C[5]	WL [6]	BT [7]	MRM	TO
0	1.70	1.68	1.71	1.69	1.71	II
0.5	1.57		1.58	1.57	1.58	II
1.0	1.40		1.41	1.40	1.41	II
1.5	1.15		1.16	1.15	1.16	II
1.75				0.95	0.96	II
1.87	0.8		0.8	0.81	0.81	II
1.9		0.76	0.76	0.77	0.76	II

1.92	0.7		0.71	0.73	0.72	II
1.95	0.65		0.65		0.65	II
1.96 2	0.62		0.62		0.62	II
1.99	0.55		0.56		0.56	I

В таблице 1 приведены значения температур только с двумя знаками после запятой для наглядности демонстрации согласования результатов в пределах 1-2 процентов. Такая погрешность является типичной для всех представленных методов. Это показывает, что метод Розе-Махты применим для исследования таких задач.

Заключение

В работе описаны результаты разработки и применения модифицированного метода Розе-Махты к модели Блуме-Капеля. Проведено сравнение критических температур, полученных с помощью исследуемого метода, с критическими температурами, полученными другими авторами и другими методами. Сделан вывод о применимости этого подхода к задаче исследования плотности состояний и критического поведения в окрестности фазового перехода для физических систем с дискретным спектром наряду с уже известными методами.

Список литературы

1. Landau, D., and K. Binder, 2021, A Guide to Monte Carlo Simulations in Statistical Physics, Cambridge University Press, Cambridge, UK.
2. Nathan Rose and Jonathan Machta. Equilibrium microcanonical annealing for first-order phase transitions. Phys. Rev. E, 100:063304, Dec 2019.
3. M. Blume, Theory of the first-order magnetic phase change in UO₂, Phys. Rev. 141 (1966) 517–524.
4. H.W. Capel, On the possibility of first-order phase transitions in Ising systems of triplet ions with zero-field splitting, Physica 32 (1966) 966–988.
5. J. C. Xavier, F. C. Alcaraz, D. Pena Lara, and J. A. Plascak, The critical behaviour of the spin-3/2 Blume-Capel model in two dimensions, Phys. Rev. B 57, 11575 1998.
6. C. J. Silva, A. A. Caparica, and J. A. Plascak, Wang-Landau Monte Carlo simulation of the Blume-Capel model, Phys. Rev. E 73, 036702, 2006
7. P. Butera and M. Pernici. The blume–capel model for spins $s=1$ and $3/2$ in dimensions $d=2$ and 3 . Physica A: Statistical Mechanics and its Applications, 507:22–66, 2018.
8. P. D. Beale, Phys. Rev. B 33, 1717 1986.

NLP-МЕТОДЫ ДЕТЕКТИРОВАНИЯ СУЩЕСТВЕННЫХ ИЗМЕНЕНИЙ В ПРОГРАММНОМ КОДЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Загороднев Д.Г.³, Кобякова А.А.¹,
Пивоваров А.С.², Попов Л.Д.²

Национальный исследовательский университет
«Высшая школа экономики»,

¹департамент прикладной математики,

²департамент компьютерной инженерии

МИЭМ НИУ ВШЭ,

³ПАО Сбербанк

Аннотация

В работе рассматриваются методы использования NLP (Natural Language Processing) механизмов для анализа изменений в версиях внедряемого программного кода. Цель работы - создание автоматического детектора изменений кода, влияющих на результат исполнения программы. Эффективность представленных механизмов продемонстрирована на реальных данных, выделенных МИЭМ НИУ ВШЭ: объектах файловой системы GitLab, содержащих историю программной реализации проектов математического моделирования.

Введение

Важность исследуемого процесса выявления существенных изменений в программном коде обусловлена необходимостью контроля качества внедряемых разрабатываемых моделей. Научные, коммерческие и другие модели и продукты программного обеспечения в крупных корпорациях подвергаются множеству доработок и обновлений разного характера в день, и использование человеческих ресурсов для тестирования и валидации каждой из них было бы неэффективным решением. В рамках данного исследования предлагается компромиссное решение: автоматизировать выделение в версиях кода изменений, способных существенно повлиять на работу моделей, и только такие изменения подвергать более глубокой проверке с привлечением специалистов.

Кроме того, представленное исследование есть новый шаг к тому, чтобы автоматизировать процесс проверки соответствия кода внедряемого программного обеспечения и технического задания. Уже сейчас проекты, частично или полностью заменяющие по функционалу аналитиков и тестировщиков (например, ChatGPT [11]), находятся в центре внимания мирового научного и инженерного сообщества [1]. Такого рода инновационные продукты уже вводятся в эксплуатацию правительством Москвы в сферы контроля безопасности и общественного порядка, как показано на Встрече лидеров Сбера при участии Сергея Собянина.

Исследуемые данные

В качестве исследуемых данных были рассмотрены программные проекты студентов и преподавателей НИУ ВШЭ, полученные из их GitLab-

репозиториях. Для детектирования изменений в проектах и ранжирования их по степени существенности была использована идея анализа разницы между состояниями проектов на различных временных интервалах, где состояние было представлено отчетной формой сохраненных внесенных изменений (далее: коммит) кода в дистрибутиве. Несмотря на потенциальную пользу анализа коммитов, существует проблема нехватки подходящего для решения задачи набора данных. Большинство обработанных размеченных наборов данных, которые можно найти, не удовлетворяют специфике нашей задачи (например, анализируются только сообщения коммита [1] или данные размечены на совершенно другие классы [2]), причем действительно информативные с профессиональной точки зрения обыкновенно отсутствуют в открытом доступе.

Исходя из этого, в работе был использован собственнo собранный и размеченный экспертно набор данных, разделенный на классы по степени значимости изменений (Рис.1). Основными требованиями к исследуемым данным были:

- Основной язык проекта - Python (механизм был бы применим и к компилируемым языкам, однако в математическом моделировании высокоуровневые языки имеют большую популярность);
- Ограниченный набор используемых библиотек (numpy, pandas, sklearn, pytorch)
- Цель исследуемого проекта - программная реализация математического моделирования разного рода систем и процессов;
- Количество произведенных коммитов превышает минимально установленный нами граничный порог.

Информация была собрана из подходящих по требованиям выделенных репозиториях GitLab, относящихся к научно-техническим проектам МИЭМ НИУ ВШЭ. Для этого был реализован скрипт по взаимодействию с GitLab API и предобработки данных для приведения их в нужный для анализа формат. Помимо самого "commit message" (текстовое поле, которое разработчик оставляет после коммита, объясняя свое изменение) и "diff" (измененный код программы в рамках рассматриваемого коммита) была извлечена информация о количестве измененных файлов, добавленных и удаленных строк и прочих информационно значимых факторах, подвергнутых дальнейшему анализу и исследованию на значимость и частотные характеристики.

commit_message	files_changed	lines_inserted	lines_deleted	diff	target
Merge branch...	9	224	180	[{'new_p	1
updated...	5	62	189	[{'new_p	0
Integrated...	4	16	7	[{'new_p	0
Integrated...	5	46	53	[{'new_p	0
Delete...	1	0	128	[{'new_p	1
Merge branch...	9	224	180	[{'new_p	1
updated...	5	62	189	[{'new_p	0
Integrated...	4	16	7	[{'new_p	0
Integrated...	5	46	53	[{'new_p	0
Delete...	1	0	128	[{'new_p	1

Рис. 1. Формат данных для обучения модели

В конечном итоге была собрана сбалансированная по классам выборка, содержащая 850 объектов с существенными и несущественными изменениями (Рис. 2):

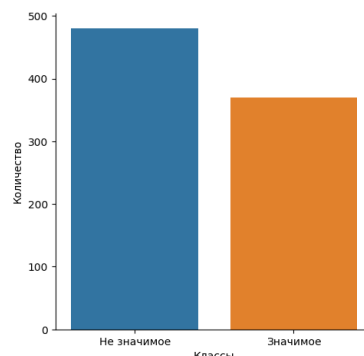


Рис. 2. Распределение классов в выборке

Классификация данных

Коммиты в системах контроля версий (например, используемый в ходе текущего исследования Git), как удобный способ отслеживания изменений, внесенных в репозитории кода, в наше время все чаще используются при решении прикладных задач [1,2]. Однако упомянутые методики либо используют только минимальную формализованную информацию из прикладываемых к коммиту сообщений, либо только из локальных изменений в разработанных объектах продуктов. Также, в этом исследовании будет учитываться их расширенная и усложненная структурно совокупность. Общая схема алгоритма решения представлена ниже на Рис. 3.



Рис. 3. Общая схема алгоритма

Подход к решению такого рода задач на основе взаимодействия с естественными языками можно разделить на 2 основных типа:

- Методы группы классического машинного обучения;
- Методы на основе сетей-трансформеров.

В представленной работе рассмотрены оба подхода и проведен анализ качества полученных моде-

лей на 3 выбранных популярных метриках: ассигасу, F1-score, Roc-Auc.

Методы классического машинного обучения

Первым этапом решения задачи построения классификатора является обработка собранных данных. Для этого случая информация из поля diff была разбита на 2 группы с добавленными и удаленными строками соответственно. Также данные были очищены от символов и стоп-слов, не обладающих информационной ценностью, например, символов переноса строки или же пустых строк.

Далее была проведена токенизация (разделение строки на составляющие, «токены»), однако большинство механизмов-токенизаторов предназначены для работы с естественными языками (разговорными), но не синтетическими (например, языками программирования). При применении подобного аппарата полученные токены не будут учитывать специфику написания программного кода. Например, название, написанное через знак нижнего подчеркивания, будет кодироваться как единое слово. Для того, чтобы избежать подобного рода проблем, был написан собственный токенизатор, учитывающий все соответствующие аспекты. После формирования токенов из программного кода и commit message было проведено удаление стоп-слов, очистка текстов от самых популярных или редких слов, а также использован стемминг (алгоритм, отбрасывающий окончания слов).

Следующий шаг включал в себя извлечение признаков (на которых базировалось обучение моделей-классификаторов). Извлечение состояло из векторизации токенов, найденных на предыдущем этапе, называемом предварительной обработкой данных. Для этого был использован метод Bag-of-words (BoW), векторизация которым приводила вид набора слов текстов в матричную форму. Кроме того, был применен метод term frequency-inverse document frequency (TF-IDF) [3]. TF-IDF — это метод поиска информации, которым производится оценка релевантности рассматриваемого слова его контексту в наборе исследуемых текстов.

Затем на финальном этапе, получив матрицу признаков после векторизации, исключив мало информативные статистические показатели, с помощью метода главных компонент (PCA) [10],

были применены несколько методов классификации на основе классического машинного обучения и рассчитаны выбранные нами метрики качества моделей. Нами были использованы последовательно следующие методы различного уровня сложности:

1. Логистическая регрессия (LogReg) – статистическая линейная модель, которая оценивает вероятность события того, что целевая переменная является линейной комбинацией одной или нескольких независимых переменных [4].

2. Случайный лес (RF), который является типом метода Bagging, а его базовым классификатором является решающее дерево. Это ансамблевый метод обучения, который строит большое количество глубоких решающих деревьев, случайным образом вы-

бирая объекты и признаки на этапе обучения, а его конечным результатом является мода классов или среднее значение отдельных деревьев.

3. Градиентный бустинг (GB) – также ансамблевая модель из деревьев решений. Основное отличие от представленного выше RF в том, что каждая базовая модель последовательно использует при обучении ошибки предшествующей версии ансамбля. Существует немалое количество различных имплементаций GB, но выбор пал на разработку от компании Яндекс - CatBoost [5].

Методы на основе сетей-трансформеров

В этом разделе основной акцент сделан на более новые и сложные методы, позволяющие учитывать не только частоту появления одного конкретного слова в тексте, но и выявлять более сложные закономерности, извлекая информацию из контекста слов и их комбинаций. Каждый из предложенных к использованию подходов основан на потенциально новой и результативной в своем прикладном применении концепции - модель «трансформер». Архитектура такого трансформера представляет из себя два блока: кодировщик и декодировщик, каждый из которых построен на использовании механизма внимания [6].

Механизм внимания, или Attention, параметризован матрицами весов запросов W_q , весов ключей W_k , весов значений W_v . Для вычисления связи, или же «внимания», входной последовательности вектора X к вектору Y , вычисляются соответствующие вектора $Q = W_q X$, $K = W_k X$, $V = W_v Y$. Эти значения используются для вычисления результата внимания по формуле (1):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (1)$$

где $softmax$ - логистическая функция.

С моделями глубокого обучения на основе механизмов трансформеров, достигающими высоких результатов в задачах NLP, Feng et al. (2020) создал проект CodeBERT [7], предварительно обученную бимодальную модель, основанную на архитектуре RoBERTa (Liu et al., 2019) (модель работы с текстовыми структурами, которая используется прежде всего для естественных разговорных языков) [8] для естественного языка и программного кода (NL-PL), чтобы зафиксировать семантическую связь между обеими модальностями с помощью механизма внимания для вывода представлений. Так, CodeBERT является специально обученной на связке NL-PL большой языковой моделью RoBERTa.

Как доказано статистически, CodeBERT результативно применяется в большом спектре задач NLP, однако при анализе программного кода модель все же работает с ним как с полноценным естественным языком, упуская информацию о табуляции, структуре и т. д. Решение проблемы нашло себя в использовании модели UniXCoder [9] от той же компании. Модель является улучшением CodeBERT и принимает на вход не только код в качестве текста, а также рас-

считывает мультимодальное содержимое, такое как комментарий к коду и абстрактное синтаксическое дерево (AST), для улучшения представления кода.

Абстрактное синтаксическое дерево используется для более детального анализа кода, преобразуя его в ориентированный ациклический граф, несущий информацию об иерархии внутри кода (Рис. 4) [9]. Такой алгоритм требует оригинального подхода к обработке данных, ведь для классических методов машинного обучения diff был разбит на композицию добавленных и удаленных строк. AST же требует на вход весь код программы, поэтому этот шаг в обработке заменен на деление diff на код до и после изменения. После построения AST алгоритм использует “one-to-one mapping” метод, с помощью которого он преобразовывает дерево в последовательную структуру, которая сохраняет всю информацию, а затем последовательность используется в качестве входных данных для улучшения представления кода (архитектура алгоритма представлена ниже на Рис. 4):

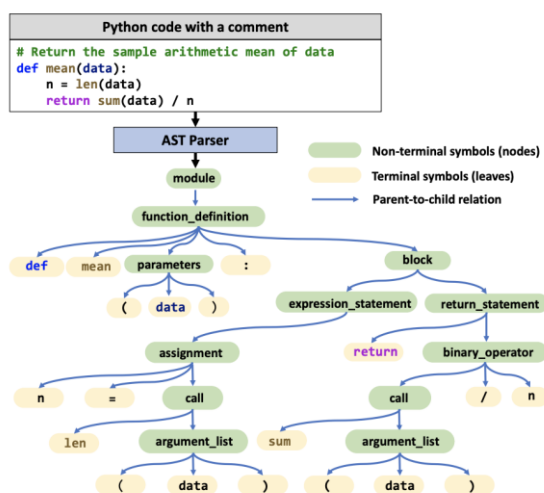


Рис. 4. Пример построения AST дерева

Таким образом, в представленном исследовании были использованы различные подходы анализа изменений со временем в коде программного обеспечения различных продуктов математического моделирования. Для реализации проекта было решено применить 2 разных типа моделей для получения численных/векторных представлений кода и commit message и использовать их в качестве новых признаков. Далее была выстроена третья модель, которая на основе всех собранных численных данных, представленных в векторной форме, выдает окончательный предикат значимости изменения версий кода.

Итого, модель RoBERTa, как было выявлено, наилучшим образом решала выстроенные задачи работы с текстовыми структурами, потому она была выбрана для анализа commit message. Также было учтено, что для наиболее корректного анализа кода, нужен особый подход при токенизации, для этого был использован токенизатор на базе CodeBERT. Далее к предобученной на естественном языке RoBERTa было добавлено два дополнительных слоя с 128 и 1 выходами соответственно и, используя

Transfer Learning (перенос обучения), были обучены эти 2 слоя модели. В качестве итогового представления были взяты выходы полученной модели из предпоследнего слоя.

В свою очередь, для анализа текста программного кода было решено использовать UniXCoder, что, прежде всего, потребовало решения проблемы различного и непредсказуемого числа измененных файлов внутри новой версии исследуемого проекта. Предложенное решение состояло в том, чтобы, используя косинусное расстояние, посчитать разницу между представлениями кода до и после, а потом просуммировать и добавить в качестве нового признака.

Для итогового классификатора выбраны те же модели, что успешно использовались в предыдущей секции (LogReg, RF, GB).

Анализ результатов

Итоговая комбинация выбранных метрик качества по всем экспериментам отражена в таблице, представленной ниже (Табл. 1). Показатели качества классификации, полученные с помощью использования нейронных сетей на основе трансформеров, являются наилучшими среди рассматриваемых методов. Как и ожидалось, лучшей связкой моделей оказалась UniXCoder и RoBERTa с финальной классификацией в виде алгоритма градиентного бустинга. Более того, после изучения примеров, на которых ошибалась данная модель, была выявлено, что каждое из этих предсказаний было сделано с довольно близким к порогу значением (модель была “не уверена”).

Таблица 1. Комбинации модели и метрики качества их применения

	Accuracy	F1_score	ROC_AUC
TF-IDF + LogReg	0,725	0,621	0,706
TF-IDF + RF	0,725	0,621	0,706
TF-IDF + GB	0,75	0,584	0,76
BoW + LogReg	0,725	0,56	0,717
BoW + RF	0,725	0,593	0,708
BoW + GB	0,75	0,583	0,76
UR + LogReg	0,825	0,588	0,781
UR + RF	0,825	0,632	0,769
UR + GB	0,9	0,778	0,891

Заключение

Итак, в ходе исследования было выявлено, что методы глубокого обучения на основе трансформеров оказались наиболее успешными в решении задачи классификации коммитов. Более того, именно разный подход к анализу сообщению коммита и программного кода помог значительно улучшить величину выбранных метрик качества решаемой задачи. Дальнейшее направления работы заключается в расширении набора данных и развитии сценариев их предобработки, с возможностью добавлять потенциально новые и информационно значимые при решении задачи признаки. Например, присутствие особых зависимых от тематики проекта ключевых слов, учет типизации изменений внедряемых версий и прочие

выделяемые показатели. Также немаловажной частью дальнейшей работы является разработка удобного пользовательского интерфейса для автоматизации процесса.

Список литературы

1. Yang Y., Ronchieri E., Canaparo M. Natural Language Processing Application on Commit Messages: a case study on HEP software //Applied Sciences. – 2022. – Т. 12. – №. 21. – С. 10773.
2. Cotroneo D. et al. Analyzing the context of bug-fixing changes in the openstack cloud computing platform //2019 IEEE 30th International Symposium on Software Reliability Engineering (ISSRE). – IEEE, 2019. – С. 334-345.
3. Zhang W., Yoshida T., Tang X. A comparative study of TF* IDF, LSI and multi-words for text classification //Expert systems with applications. – 2011. – Т. 38. – №. 3. – С. 2758-2765.
4. Dönmez, P. Introduction to Machine Learning. *Nat. Lang. Eng.* 2013, 19, 285–288.
5. Dorogush A. V., Ershov V., Gulin A. CatBoost: gradient boosting with categorical features support //arXiv preprint arXiv:1810.11363. – 2018.
6. Vaswani A. et al. Attention is all you need //Advances in neural information processing systems. – 2017. – Т. 30.
7. Feng Z. et al. Codebert: A pre-trained model for programming and natural languages //arXiv preprint arXiv:2002.08155. – 2020.
8. Liu Y. et al. Roberta: A robustly optimized bert pretraining approach //arXiv preprint arXiv:1907.11692. – 2019.
9. Guo D. et al. Unixcoder: Unified cross-modal pre-training for code representation //arXiv preprint arXiv:2203.03850. – 2022.
10. Ian T. Jolliffe and Jorge Cadima: Principal component analysis: a review and recent developments //https://doi.org/10.1098/rsta.2015.0202 - 2016
11. Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, Yi Zhang: Sparks of Artificial General Intelligence: Early experiments with GPT-4 // https://arxiv.org/abs/2303.12712 - 2023

ОСЦИЛЛИРУЮЩЕЕ РАСЩЕПЛЕНИЕ НИЖНИХ ЭНЕРГЕТИЧЕСКИХ УРОВНЕЙ КВАДРАТИЧНОГО ОПЕРАТОРА НА АЛГЕБРЕ $SU(1, 1)$

Румянцева С.В.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

Данная работа посвящена математическому исследованию квазиклассических туннельных асимп-

тотик нижних уровней в дискретном спектре операторов на алгебрах Ли. Хорошо известно, что в задачах о спектре оператора Шредингера с потенциалом типа двойной ямы асимптотика туннельного расщепления энергий монотонно убывает по формуле Ландау-Лифшица. С другой стороны, при исследовании мономолекулярных магнитов (Fe_8 , Mg_{12}) было показано, что для операторов на алгебре Ли $su(2)$, асимптотика расщепления пары близких энергий имеет нестандартный вид: в асимптотике возникает дополнительный осциллирующий множитель. В данной работе мы исследуем появление осциллирующего туннельного эффекта в задаче на алгебре Ли $su(1,1)$. Применяя когерентное преобразование и метод комплексных ВКБ асимптотик, мы доказали теорему о квазиклассической асимптотике туннельного расщепления нижних энергетических уровней для квадратичного оператора.

Введение

Интерес к изучению туннельного эффекта вызван тем, что туннелирование играет ключевую роль в некоторых квантовых системах. Например, в трополоне возникают процессы туннелирования атомов водорода и тяжелых атомов [1] или в мономолекулярных магнитах возникает спиновое туннелирование [2].

Базовой моделью исследования туннельного эффекта является оператор Шредингера с потенциалом в виде двойной ямы [4, 5]. Для одномерного случая асимптотическая формула для величины расщепления имеет вид [6]:

$$\Delta E = \frac{\hbar\omega}{\pi} \exp\left(-\frac{1}{\hbar} \int \sqrt{V(x) - E} dx\right) (1 + O(\hbar)) \quad (1)$$

При исследовании мономолекулярных магнитов возникает квантово-механическая модель с оператором на алгебре Ли $su(2)$ [2, 3]. Тогда асимптотика туннельного расщепления энергий отличается от классической формулы Ландау-Лифшица (1) дополнительным осциллирующим множителем [3]. Появление осциллирующего эффекта в расщеплении энергии вызывает большой интерес (см. в [7]).

Данная работа посвящена исследованию осциллирующего эффекта туннелирования для нижних энергетических уровней оператора на алгебре Ли $su(1,1)$. Задачи, связанные с нижними энергетическими уровнями, представляют большой интерес в квантовом моделировании [8, 9]. Это обусловлено тем, что нижние уровни энергии оказываются более заселенными при низких температурах.

Мы рассматриваем симметричный квадратичный оператор:

$$\hat{H} = \frac{1}{2} \hat{B}^2 + \frac{1}{2} \hat{B}^{*2} + \alpha(\hat{A} - \beta)^2, \quad (2)$$

где операторы $\hat{A}, \hat{B}, \hat{B}^*$ являются элементами алгебры $su(1, 1)$. Траектории классического движения для гамильтониана на единичном диске при $\alpha = 6\sqrt{5}$ и $\beta = 2\sqrt{5}$ имеют вид двойной ямы (Рис. 1).

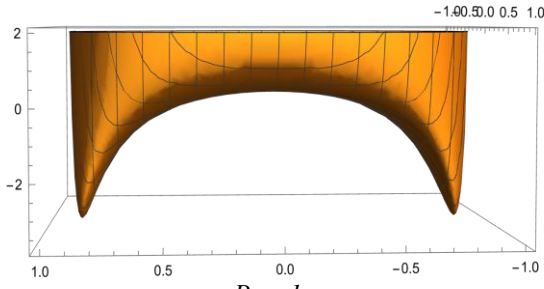


Рис. 1

Ранее в работе [10] для оператора (2) была доказана теорема об асимптотике туннельного расщепления верхних энергетических уровней:

$$\Delta E_n = (-1)^n \frac{2\hbar\omega}{\pi} \exp\left\{-\frac{i}{2\hbar} \oint \theta^c\right\} \sin\left(\frac{i\pi}{\hbar} \text{Res}_{z=\xi_0} \theta^c|_{\Lambda^c} + O(\hbar)\right) (1 + O(\hbar)), \quad (3)$$

где номер n порядка $1/\hbar$. Заметим, что величина расщепления (3) отличается от классической формулы Ландау-Лифшица (1) наличием осциллирующего множителя, как и для оператора на $su(2)$.

В настоящей работе мы исследуем вопрос о сохранении осциллирующего эффекта туннелирования для нижних энергетических уровней оператора (2) (при конечных n).

Для решения задачи о построении асимптотики разности пары близких энергий мы согласовываем асимптотику собственных функций оператора (2) в окрестностях двукратных точек поворота (см. вывод формулы (5) в следующем разделе). Особенность исследования нижних уровней заключается в необходимости рассматривать не только ВКБ асимптотики решений, но и асимптотики в терминах специальных функций параболического цилиндра [11, 12].

Квазиклассическая асимптотика туннелирования нижних энергетических уровней

Пусть $\alpha > 1$ и $|\alpha|2\beta - 1| < 1$. Исследуем спектральную задачу для оператора (2) вблизи минимального значения энергии:

$$E = -\frac{\alpha(4\beta^2 - 1) + 1}{4(\alpha - 1)} + \hbar E^1.$$

Ранее в [10] было показано, что после применения когерентного преобразования [13] и преобразования Лиувилля спектральная задача для оператора (2) переходит в дифференциальное уравнение Шредингера в пространстве антиголоморфных представлений:

$$\hbar^2 u''(\bar{z}) - \left(Q(\bar{z}) - \frac{2E}{\bar{z}^4 - 2\alpha\bar{z}^2 + 1} + O(\hbar^2)\right) u(\bar{z}) = 0, \quad (4)$$

где α, β – действительные постоянные, \hbar – малый положительный параметр, $Q(\bar{z})$ имеет вид:

$$Q(\bar{z}) = \frac{z^4 \alpha(1 + 2\beta)^2 - 2z^2 + \alpha(1 - 2\beta)^2}{2(\bar{z}^4 - 2\alpha\bar{z}^2 + 1)^2}.$$

Фазовым пространством дифференциального уравнения (4) является круг единичного радиуса. Точки спектра оператора (2) определяются из условия аналитичности решений уравнения (4) в круге $|z| < 1$ [10].

Придерживаясь подхода работы [10], мы выводим асимптотики туннельного расщепления энергий исследуемого оператора (2), основываясь на построении матрицы обхода вокруг особой точки $\xi_0 = \sqrt{\alpha - \sqrt{\alpha^2 - 1}}$ уравнения (4). Для этого необходимо согласовать асимптотики решений, полученных с помощью комплексного метода ВКБ, и асимптотики решений, записанных через функции параболического цилиндра.

Пусть z_0 – положительная двукратная точка поворота, в которой $Q(\bar{z}) = 0$. Пусть выполнено правило квантования:

$$E^1 = \sqrt{\frac{Q''(z_0)}{2}} (-z_0^4 + 2\alpha z_0^2 - 1)(n + 1/2), \quad n \in \mathbb{Z}_+$$

Получаем, что асимптотика туннельного расщепления энергий оператора (2) в случае нижних уровней энергий имеет вид (подробное доказательство будет опубликовано в отдельной работе):

$$\Delta E_n = \sqrt{\frac{Q''(z_0)}{2}} (z_0^4 - 2\alpha z_0^2 + 1) \frac{(-1)^{n+1} 4\sqrt{\hbar}}{n! \sqrt{\pi}} \left(\frac{2}{\hbar}\right)^n J^{2n+1} \exp\left\{-\frac{1}{\hbar} \int_{-z_0}^{z_0} \sqrt{Q(\bar{z})} dz\right\} \sin\left(\frac{2\pi}{\hbar} \text{Res}_{z=\xi_0} \sqrt{Q(\bar{z})} + o(1)\right) (1 + o(1)), \quad (5)$$

где коэффициент J в амплитуде выражается в виде:

$$J = \sqrt{\frac{Q''(z_0)}{2}} z_0 \exp\left(-\int_0^{z_0} \sqrt{\frac{Q''(z_0)}{2} \frac{-z^4 + 2\alpha z^2 - 1}{Q(t) - z_0^4 + 2\alpha z_0^2 - 1}} - \frac{1}{t - z_0} dt\right).$$

Отметим, что асимптотика расщепления для нижних энергетических уровней (5) переходит в асимптотику расщепления для высоких энергетических уровней (3) при предельном переходе $n \rightarrow \infty$.

Заключение

В данной работе мы рассматривали задачу о построении асимптотики туннельного расщепления энергий оператора (2), заданного на неприводимом представлении алгебры Ли $su(1,1)$, для нижних энергетических уровней энергии. В отличие от (3),

где $n \rightarrow \infty$ порядка $1/\hbar$, асимптотика (5) построена для конечных n .

Асимптотические формулы (3) и (5) показывают, что в данной задаче туннельное расщепление спектра не только экспоненциально убывает, но и быстро осциллирует при $\hbar \rightarrow 0$, как и для алгебры $su(2)$. Это нетипичный эффект, который невозможен в случае стандартного одномерного уравнения Шредингера. Данный эффект объясняется туннелированием по двум гомотопически различным инстантам на комплексном лагранжевом многообразии и возникновении интерференции при туннелировании.

В дальнейшем планируется представить формулу (5) в квазиклассическом виде, введя лагранжево многообразие и соответствующие физические величины (частоту, действие, инстантон), как в работе [10]. Тогда можно будет заметить, что асимптотика расщепления энергии выражается через комплексификацию фазового пространства и лагранжево многообразия. Отметим, что обычное лагранжево многообразие не дает представление о том, как проявляется квантовое туннелирование.

Список литературы

1. Redington R. L. H atom and heavy atom tunneling processes in tropolone //The Journal of Chemical Physics. – 2000. – Т. 113. – №. 6. – С. 2319-2335.
2. Garg A. Quenched spin tunneling and diabolical points in magnetic molecules. I. Symmetric configurations // Phys. Rev. B – 2001. – V. 64. – №. 9. – P. 094413.
3. Garg A. Quenched spin tunneling and diabolical points in magnetic molecules. II. Asymmetric configurations //Phys. Rev. B. – 2001. – V. 64. – №. 9. – P. 094414.
4. Song D. Y. Tunneling and energy splitting in an asymmetric double-well potential //Annals of Physics. – 2008. – Т. 323. – №. 12. – С. 2991-2999.
5. Miller Jr S. C., Good Jr R. H. A WKB-type approximation to the Schrödinger equation //Physical Review. – 1953. – Т. 91. – №. 1. – С. 174.
6. Ландау, Л. Д. Квантовая механика. Нерелятивистская теория / Л. Д. Ландау, Е. М. Лифшиц // Теоретическая физика. — Т. 3. — Издание 1-е. — Л.: Гос. изд-во РСФСР, 1948. — 567 с.
7. Брюнинг Й., Доброхотов С. Ю., Некрасов Р. В. Расщепление нижних энергетических уровней в квантовой двойной яме в магнитном поле и туннелирование волновых пакетов в нанопроводах //Теоретическая и математическая физика. – 2013. – Т. 175. – №. 2. – С. 206-225.
8. Bell R. P. The tunnel effect in chemistry. – Springer, 2013.
9. Гольданский В. И., Трахтенберг Л. И., Флёргов В. Н. Туннельные явления в химической физике. – 1986
10. Vybornyi E. V., Rumyantseva S. V. Semiclassical Asymptotics of Oscillating Tunneling for a Quadratic Hamiltonian on the Algebra $su(1, 1)$ //Math. Notes – 2022. – V. 112. – №. 5. – P. 665-681.
11. Федорюк М. В. Асимптотические методы для линейных обыкновенных дифференциальных уравнений. // URSS, 2009.
12. Славянов С. Ю. Асимптотика сингулярных задач Штурма–Лиувилля по большому параметру в случае близких точек перехода //Дифференциальные уравнения. – 1969. – Т. 5. – No. 2. – С. 313-325.
13. M. Karasev, E. Novikova Non-Lie permutation representations, coherent states, and quantum embedding // Translations of the American Mathematical Society-Series 2 187, 1-202.

ДИФФЕРЕНЦИАЛЬНАЯ ИГРА ПЕРЕХВАТА И ЗАЩИТЫ ЦЕЛИ

Семион А.А.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ,
Федеральный исследовательский центр
«Информатика и управление» РАН*

Аннотация

В данной работе исследуется дифференциальная игра преследования с несколькими игроками: нападающим, целью которого является сближение с некоторым объектом, и защитниками, целью которых является перехват нападающего. Задача управления несколькими игроками рассматривается, как игра с нулевой суммой. Приводятся законы управления игроками. Проверка законов управления произведена с помощью математического моделирования.

Введение

Теория дифференциальных игр зародилась в середине 50х и связана с работами R.P. Isaacs [1], J.V. Breackwell [2], Л.С. Понтрягина [3, 4] и других зарубежных и российских ученых. В работе [1] впервые было рассмотрены приложения теории дифференциальных игр к различным военным задачам.

С конца 70-х годов прошлого века появилась, самостоятельная часть прикладной теории дифференциальных игр, в которой рассматриваются задачи преследования, задачи уклонения, задачи защиты цели.

В данной работе рассмотрена задача перехвата некоторого нападающего, в качестве которого может выступать ракета или беспилотный летательный аппарат, другими аппаратами - защитниками.

Для оценки действий нападающего и защитников вводится общий функционал качества для игры с нулевой суммой. Необходимо построить стратегию защиты цели с несколькими игроками, имеющими равную полную информацию об игре. Один из игроков должен достичь некоторой цели и уклониться от остальных игроков, задачей которых является его перехват.

Похожие постановки задачи возникают при подборе стратегий противоракетной обороны. Например, в статье [5] рассматривается перехват сверхзву-

ковой ракеты ракетами ПРО, как дифференциальная игра. Однако при синтезе законов управления игроками не предполагается какая-либо настройка приоритетов уклонения и достижения цели.

В качестве основы в данной работе будет взята статья [6], результаты которой будут модифицированы путем добавления в игру защищаемого объекта и коэффициентов, определяющих приоритеты уклонения и достижения цели нападающим.

Приведены результаты моделирования. Исследуемая игра с нулевой суммой может быть полезна для планирования стратегий на конечной стадии преследования, в котором участвуют несколько преследующих и один нападающий.

Постановка задачи

Рассматривается дифференциальная игра, в которой участвуют несколько игроков. В пространстве находится защищаемый объект, описываемый координатами g .

В это пространство попадает нападающий игрок и несколько защитников, задача которых заключается в том, чтобы перехватить нападающего. Задачей нападающего является приблизиться на некоторое расстояние к защищаемому объекту, уклонившись от преследователей. Динамика каждого объекта описывается стационарной линейной системой дифференциальных уравнений со скалярным управлением. В данной работе предполагается, что каждый игрок наблюдает непосредственное положение остальных. Игра заканчивается, если один из защитников приблизился к нападающему на расстояние ε , если нападающий приблизился к цели на расстояние ε или по истечению времени игры t_f .

Предположим, что дифференциальная игра «задача преследования» имеет место в n -мерном Евклидовом пространстве, где положения игроков могут быть записаны через вектора для нападающего, $y = [y_1, y_2, \dots, y_m]^T$ и $x_j = [x_{j1}, x_{j2}, \dots, x_{jm}]^T$, где $j = 1, 2, \dots, n$, для защитников соответственно. Вектор положения защищаемого объекта будет обозначаться как $g = [g_1, g_2, \dots, g_m]^T$.

Введем вектор $z_{yj}(t) = x_j(t) - y(t)$,

обозначающий расстояние между нападающим и j -тым защитником. В более компактном виде

$$z_y(t) = x(t) - \mathbf{1}_n \otimes y(t),$$

где $\mathbf{1}_n$ - вектор столбец размера $n \times 1$ с элементами, равными 1, и \otimes обозначает произведение Кронекера.

Дополнительно введем вектор, показывающий расстояние между нападающим $y(t)$ и защищаемым объектом $g(t)$:

$$z_g(t) = \mathbf{1}_n \otimes (y(t) - g(t)).$$

Введем вектор z являющийся взвешенной суммой векторов z_y и z_g .

$$z = h_y z_y + h_g z_g.$$

Коэффициенты h_y и h_g выбираются в соответствии с целями, стоящими перед нападающим по уклонению или достижению цели соответственно.

В работе мы будем искать управления непосредственно скоростями каждого из игроков:

$$u_p(t) = \frac{d}{dt} x(t) - \text{вектор скоростей защитников},$$

$$u_e(t) = \frac{d}{dt} y(t) - \text{вектор скоростей нападающего}.$$

В [6] было показано, что для задачи преследования возможно ввести общий для нападающих и защитников функционал и рассматривать задачу управления игроками как игру с нулевой суммой:

$$J_\Sigma(z(\cdot), u_p(\cdot), u_e(\cdot)) = \frac{1}{2} z^T(t_f) F z(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left\{ z^T(t) Q z(t) + u_p^T(t) R u_p(t) - (\mathbf{1}_n \otimes u_e(t))^T P (\mathbf{1}_n \otimes u_e(t)) \right\} dt,$$

где $F = [k_{pf} + k_{ef}] I_n$ - матрица штрафа по положениям игроков характеризующая конечную стоимость игры,

$Q = [q_p + q_e] I_n$ - матрица штрафа по положениям игроков во время игры.,

$R = r_p I_n$, $P = r_e I_n$, - матрицы штрафа за излишний расход энергии защитниками и нападающим,

k_{pf} , k_{ef} , q_p , q_e , r_p , r_e положительны, I_n - единичная матрица, размерностью $n \times n$.

В [6] показано, что управления для игроков представляются в виде:

$$u_p(t) = -\frac{1}{r_p} K(t) z(t),$$

$$u_e(t) = -\frac{1}{nr_e} (\mathbf{1}_n^T \otimes I_m) K(t) z(t),$$

где

$$\frac{d}{dt} K(t) = -K(t) \left[-\frac{1}{r_p} I_n + \frac{1}{nr_e} (\mathbf{1}_n \otimes \mathbf{1}_n^T \otimes I_m) \right] K(t) - [q_p + q_e] I_n,$$

$$K(t) = [k_{pf} + k_{ef}] I_n.$$

Также в [6] доказывается, что игра имеет цену при $r_p < nr_e$. Это ограничение можно рассматривать, как ограничение на отношение скоростей игроков.

Проверка работы данных стратегий проведена путем математического моделирования.

Моделирование

Для проверки работы синтезированного управления рассматривается игра с 4мя игроками на плоскости.

Начальные положения игроков следующие: $x_1(t_0) = [4, 1]^T$, $x_2(t_0) = [3, 0]^T$, $x_3(t_0) = [3, -1]^T$, $y(t_0) = [0, 3]^T$. Координаты защищаемого объекта: $g = [-1, 2]^T$. Конечное время игры: 5с.

Проведем математическое моделирование игры. Траектории игроков показаны на рис. 1.

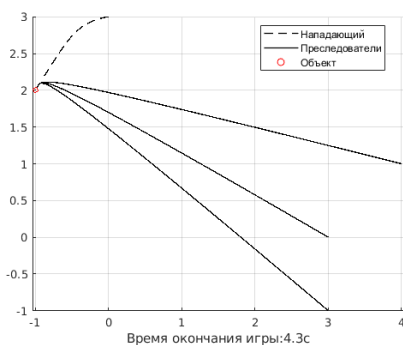


Рис. 1. Трактории игроков

Игра закончена в связи с перехватом нападающего.

По траекториям можно видеть, как нападающий старается не просто двигаться напрямую к объекту, но также уклоняется от преследователей. Таким образом, преследователям пришлось проделать больший путь. Однако, скоростей преследователей хватило для перехвата нападающего.

Заключение

В работе рассмотрена задача перехвата некоторого аппарата, стремящегося сблизиться с защищаемым объектом, несколькими перехватчиками. Задача рассмотрена как игра с нулевой суммой. Предложен подход с введением взвешенной суммы в вектор состояния игры, для выбора приоритетной задачи (уклонения от перехвата или достижения объекта).

Приведены стратегии для всех игроков, проведено математическое моделирование поведения игроков, использующих предложенные стратегии.

Моделирование показало работоспособность предложенного подхода.

В дальнейшем исследовании стоит провести анализ эффективности управления с учетом начальных положений игроков и защищаемого объекта.

Работа выполнена при финансовой поддержке РНФ (проект № 21–11–00202).

Список литературы

1. Айзекс Р. Дифференциальные игры. М.: Мир, 1967. 480 с
2. Breackwell J.V., and Merz A.W. Toward a Complete Solution of the Homicidal Chauffeur Game, Proceedings of the 1st International Conference on the Theory and Applications of Differential Games, Amherst, Massachusetts, 1969
3. Понtryгин Л. С. О линейных дифференциальных играх. 1. ДАН, 1967, т. 174, № 6, с. 1278—1280.
4. Понtryгин Л. С. О линейных дифференциальных играх. 2. ДАН, 1967, т. 175, № 4, с. 764—766.
5. Liang H., Li Z., Wu J., Zheng Y., Chu H., Wang J. Optimal Guidance Laws for a Hypersonic Multiplayer Pursuit-Evasion Game Based on a Differential Game Strategy // Aerospace. 2022. Т. 9. № 2. С. 97.
6. Афанасьев В.Н., Семион А.А. Дифференциальные игры преследования с несколькими преследователями и одним уклоняющимся // Проблемы управления. 2021. №1., с. 24–35.

СВОЙСТВА АГРЕГИРОВАННОЙ КВАЗИГАЗОДИНАМИЧЕСКОЙ СИСТЕМЫ УРАВНЕНИЙ МНОГОКОМПОНЕНТНОЙ ГАЗОВОЙ СМЕСИ

Федченко А.С.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент математики,
факультет экономических наук

Аннотация

В данной работе для агрегированной системы регуляризованных (квазигазодинамических) уравнений односкоростной и однотемпературной многокомпонентной газовой смеси получено уравнение баланса энтропии с неотрицательным производством энтропии при наличии потоков диффузии. Для системы, линеаризованной на постоянном решении, выведены существование, L^2 – диссипативность и единственность слабых решений начально-краевой задачи. Для самой квазигазодинамической системы получены параболичность по Петровскому и локальная по времени классическая однозначная разрешимость задачи Коши.

Введение

Регуляризованные, или квазигазодинамические (КГД), системы уравнений давно зарекомендовали себя в компьютерном моделировании задач динамики однокомпонентного газа [1]. Поскольку многие газы и жидкости, встречающиеся в природе и эксплуатируемые в промышленности, представляют собой смесь нескольких компонент, исследователей также интересует численное моделирование течений многокомпонентных, в том числе бинарных, смесей газов и жидкостей [1]–[3].

В данной работе выполняется исследование агрегированной системы КГД уравнений односкоростной и однотемпературной многокомпонентной смеси при наличии потоков диффузии (см. в этой связи [4,5]).

Агрегированная КГД система уравнений многокомпонентной смеси сжимаемых газов

Агрегированная КГД система уравнений односкоростной и однотемпературной многокомпонентной смеси при наличии потоков диффузии включает в себя уравнения баланса массы компонент, суммарного импульса и суммарной полной энергии

$$\partial_t \rho_\alpha + \operatorname{div} [\rho_\alpha (\mathbf{u} - \mathbf{w}_\alpha) + \mathbf{d}_\alpha] = 0, \quad (1)$$

$$\begin{aligned} \partial_t (\rho \mathbf{u}) + \operatorname{div} [\rho (\mathbf{u} - \mathbf{w}) \otimes \mathbf{u}] + \nabla p = \\ = \operatorname{div} \Pi + [\rho - \tau \operatorname{div} (\rho \mathbf{u})] \mathbf{f}, \end{aligned} \quad (2)$$

$$\begin{aligned} \partial_t E + \operatorname{div} \left[\frac{1}{2} \rho |\mathbf{u}|^2 (\mathbf{u} - \mathbf{w}) + \langle \rho_\alpha h_\alpha (\mathbf{u} - \mathbf{w}_\alpha) \rangle \right] = \\ = \operatorname{div} (-\mathbf{q} + \Pi \mathbf{u}) + \rho (\mathbf{u} - \mathbf{w}) \cdot \mathbf{f} + Q \end{aligned} \quad (3)$$

с $\alpha = \overline{1, K}$. В системе ∂_t и ∂_i – частные производные по времени t и по пространственным координатам x_i . Операторы дивергенции div и градиента ∇ берутся по $\mathbf{x} = (x_1, \dots, x_n)$, $n = 1, 2, 3$. Для операции

суммирования по индексу α вводится обозначение $\langle \cdot \rangle$, символы \otimes и \cdot служат для обозначения тензорного и скалярного произведения векторов соответственно. Дивергенция тензора берется по первому индексу тензора.

В данной системе плотности компонент $\rho_\alpha > 0$, $\alpha = \overline{1, K}$, скорость $\mathbf{u} = (u_1, \dots, u_n)$ и температура $\theta > 0$ смеси являются основными искомыми функциями. Искомые функции зависят от x и t . Рассматривается модель, в которой компоненты смеси представляют собой идеальные политропные газы. Уравнения состояния имеют вид:

$$p_\alpha = R_\alpha \rho_\alpha \theta, \quad \varepsilon_\alpha = c_{V\alpha} \theta, \quad \alpha = \overline{1, K},$$

где p_α и ε_α — давление и удельная внутренняя энергия компоненты α , с постоянными $\gamma_\alpha > 1$, $R_\alpha = (\gamma_\alpha - 1)c_{V\alpha} > 0$ и $c_{V\alpha} > 0$. Также вводится $h_\alpha = \varepsilon_\alpha + (p_\alpha/\rho_\alpha) = c_{p\alpha}\theta$ — удельная энтальпия компоненты α , а $c_{V\alpha}$ и $c_{p\alpha} = c_{V\alpha} + R_\alpha$ — ее удельные теплоемкости.

Суммарная плотность, давление, удельная внутренняя энергия и полная энергия смеси вводятся следующим образом:

$$\rho = \langle \rho_\alpha \rangle, \quad p = \langle p_\alpha \rangle = R\rho\theta, \\ \varepsilon = \langle c_\alpha \varepsilon_\alpha \rangle = c_V \theta, \quad E = \frac{1}{2} \rho |\mathbf{u}|^2 + \rho \varepsilon,$$

с функциями $R := \langle c_\alpha R_\alpha \rangle$, $c_V = \langle c_\alpha c_{V\alpha} \rangle$, а не постоянными, как в однокомпонентном случае, $c_\alpha := \frac{\rho_\alpha}{\rho}$ — массовые концентрации компонент смеси. В уравнениях также участвуют регуляризующие скорости компоненты α и суммарные скорости вида

$$\mathbf{w}_\alpha = \frac{\tau}{\rho_\alpha} [\operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p_\alpha - \rho_\alpha \mathbf{f}], \\ \widehat{\mathbf{w}}_\alpha = \tau \left[(\mathbf{u} \cdot \nabla) \mathbf{u} + \frac{1}{\rho_\alpha} \nabla p_\alpha - \mathbf{f} \right], \\ \mathbf{w} := \langle c_\alpha \mathbf{w}_\alpha \rangle, \quad \widehat{\mathbf{w}} := \langle c_\alpha \widehat{\mathbf{w}}_\alpha \rangle,$$

где $\tau = \tau(\rho, \mathbf{u}, \theta) > 0$ — параметр регуляризации, а $\rho := (\rho_1, \dots, \rho_K)$.

В уравнениях (2), (3) тензор вязких напряжений Π имеет вид $\Pi = \Pi^{NS} + \Pi^\tau$. Поток тепла имеет вид $\mathbf{q} = \mathbf{q}^F + \mathbf{q}^d + \mathbf{q}^\tau$. Тензор вязкости Навье-Стокса Π^{NS} и поток тепла Фурье задаются стандартными формулами

$$\Pi^{NS} = \mu \left[\nabla \mathbf{u} + (\nabla \mathbf{u})^T - \frac{2}{3} (\operatorname{div} \mathbf{u}) \mathbb{I} \right] + \lambda (\operatorname{div} \mathbf{u}) \mathbb{I}, \\ -\mathbf{q}^F = \kappa \nabla \theta,$$

где \mathbb{I} — единичный тензор порядка n , $\mu > 0$, $\lambda \geq 0$ и $\kappa > 0$ — суммарные коэффициенты динамической и объемной вязкости и теплопроводности. Указанные коэффициенты могут зависеть от искоемых функций $(\rho, \mathbf{u}, \theta)$, а также от $\nabla \mathbf{u} = \{\partial_i u_j\}_{i,j=1}^n$.

Регуляризующие тензор вязкости и поток тепла имеют вид

$$\Pi^\tau = \rho \mathbf{u} \otimes \widehat{\mathbf{w}} + \tau [\mathbf{u} \cdot \nabla p + \langle \gamma_\alpha p_\alpha \rangle \operatorname{div} \mathbf{u} -$$

$$-\langle (\gamma_\alpha - 1) Q_\alpha \rangle] \mathbb{I}, \\ -\mathbf{q}^\tau = \tau \{ [c_V \rho \nabla \theta - \theta \nabla (R\rho)] \cdot \mathbf{u} - \langle Q_\alpha \rangle \} \mathbf{u}.$$

Величины $\mathbf{f}(x, t)$ и $Q_\alpha \geq 0$ — плотность массовой силы и мощности тепловых источников соответственно. Они считаются заданными функциями.

Для бинарной смеси ($K = 2$) при выполнении условий $\mathbf{d}_\alpha = 0$ и $\mathbf{q}^d = 0$ эти уравнения были получены в [5] агрегированием по α КГД уравнений смеси из [4].

Модель, рассматриваемая в данной работе, представляет собой регуляризованную систему уравнений Навье-Стокса односкоростной и однотемпературной смеси вязких теплопроводных сжимаемых газов. Поток диффузии и дополнительный поток тепла введем, как и в [6], формулами

$$-\mathbf{d}_\alpha := d_0 \left[\sum_{\beta: \beta \neq \alpha} (G_\alpha - G_\beta) + b_\alpha \nabla \theta \right], \quad \alpha = \overline{1, K}, \\ \mathbf{q}^d = \langle (G_\alpha + K^{-1} b_\alpha \theta) \mathbf{d}_\alpha \rangle,$$

где $G_\alpha = \varepsilon_\alpha - s_\alpha \theta + \frac{p_\alpha}{\rho_\alpha} = (c_{p\alpha} - s_\alpha) \theta$ — потенциал Гиббса и $s_\alpha = s_{\alpha 0} - R_\alpha \ln \rho_\alpha + c_{V\alpha} \ln \theta$ — удельная энтропия компоненты α , с $s_{\alpha 0} = \text{const}$. Величины $d_0 \geq 0$ и b_α здесь не конкретизируются, предполагается, что $\langle b_\alpha \rangle = 0$, следовательно, и $\langle \mathbf{d}_\alpha \rangle = 0$.

Уравнение баланса энтропии

Введем $S := \langle c_\alpha s_\alpha \rangle$ — суммарную удельную энтропию.

Теорема 1. Пусть $d_0 > 0$. При наличии потоков диффузии для рассматриваемой модели односкоростной и однотемпературной многокомпонентной смеси верно следующее регуляризованное уравнение баланса энтропии

$$\partial_t (\rho S) + \operatorname{div} [\langle \rho_\alpha s_\alpha (\mathbf{u} - \mathbf{w}_\alpha) \rangle + \frac{1}{K} \langle b_\alpha \mathbf{d}_\alpha \rangle + \\ + \frac{1}{\theta} (\mathbf{q}^F + \mathbf{q}^\tau)] = \mathcal{P}^{NS} + \langle \mathcal{P}^\tau \rangle$$

с неотрицательным производством энтропии $\mathcal{P}^{NS} + \langle \mathcal{P}^\tau \rangle$, где $\mathcal{P}^{NS} \geq 0$ — производство энтропии по Навье-Стоксу, $\langle \mathcal{P}^\tau \rangle$ — регуляризующее слагаемое производства энтропии, $\mathcal{P}_\alpha^\tau \geq 0$ при условии $\tau(\gamma_\alpha - 1)Q_\alpha \leq 4p_\alpha$, $\alpha = \overline{1, K}$.

Линеаризация КГД системы уравнений односкоростной и однотемпературной многокомпонентной смеси и ее диссипативность

При $Q_1 = \dots = Q_\alpha = 0$, $\mathbf{f} = 0$ КГД система (1)-(3) имеет постоянные решения

$$\mathbf{z}_0 = (\rho_{10}, \dots, \rho_{K0}, \mathbf{u}_0, \theta_0), \\ \text{где } \rho_{\alpha 0} > 0, \alpha = \overline{1, K}, \mathbf{u}_0 = (u_{01}, \dots, u_{0n}), \theta_0 > 0.$$

Представим решение в виде

$$\rho_\alpha = \rho_{\alpha 0} + \rho_{\alpha*} \tilde{\rho}_\alpha, \quad \mathbf{u} = \mathbf{u}_0 + u_* \tilde{\mathbf{u}}, \quad \theta = \theta_0 + \theta_* \tilde{\theta},$$

где $\rho_{\alpha*} > 0, \alpha = \overline{1, K}, u_* \neq 0, \theta_* > 0$ являются безразмеривающими параметрами, $\tilde{\mathbf{z}} := (\tilde{\rho}, \tilde{\mathbf{u}}, \tilde{\theta})$ – вектор безразмерных возмущений.

Следуя [6], линейаризацию удобнее выполнять, предварительно производя редукцию уравнений с точностью $O(|\nabla \mathbf{z}|^2)$, где $\mathbf{z} = (\rho, \mathbf{u}, \theta)$. Введем также фоновое нормированное решение $(\hat{\rho}_{10}, \dots, \hat{\rho}_{K0}, \hat{\mathbf{u}}_0, \hat{\theta}_0) := D_*^{-1} \mathbf{z}_0$ с $\hat{\mathbf{u}}_0 = (\hat{u}_{10}, \dots, \hat{u}_{n0})$ с матрицей $D_* := \text{diag}(\rho_{1*}, \dots, \rho_{K*}, u_*, \dots, u_*, \theta_*)$. Введем также фоновые значения величин ρ, c_α, R, c_V и $R_\alpha \gamma_\alpha, R_\alpha^2$:

$$\rho_0 := \langle \rho_{\alpha 0} \rangle, c_{\alpha 0} := \frac{\rho_{\alpha 0}}{\rho_0}, R_0 := \langle c_{\alpha 0} R_\alpha \rangle, c_{V0} = \langle c_{\alpha 0} c_{V\alpha} \rangle, \\ (R\gamma)_0 = \langle c_{\alpha 0} R_\alpha \gamma_\alpha \rangle, (R^2)_0 = \langle c_{\alpha 0} R_\alpha^2 \rangle.$$

Далее выполняется подстановка решения в форме $\mathbf{z} = \mathbf{z}_0 + D_* \tilde{\mathbf{z}}$ в редуцированную систему. Затем в полученной системе отбрасываются слагаемые 2-го порядка малости относительно вектор-функции $\tilde{\mathbf{z}}$ и ее производных 1-го и 2-го порядка. В итоге получается линейаризованная система уравнений

$$\begin{aligned} & \partial_t \tilde{\rho}_\alpha + u_* (\hat{\mathbf{u}}_0 \cdot \nabla \tilde{\rho}_\alpha + \hat{\rho}_{\alpha 0} \text{div} \tilde{\mathbf{u}}) = \\ & = \tau_0 u_*^2 \left[\frac{R_\alpha \theta_0}{u_*^2} \Delta \tilde{\rho}_\alpha + (\hat{\mathbf{u}}_0 \cdot \nabla)^2 \tilde{\rho}_\alpha + \right. \\ & \left. + 2\hat{\rho}_{\alpha 0} (\hat{\mathbf{u}}_0 \cdot \nabla) \text{div} \tilde{\mathbf{u}} + \frac{R_\alpha \hat{\rho}_{\alpha 0} \theta_*}{u_*^2} \Delta \tilde{\theta} \right], \alpha = \overline{1, K}, \\ & \partial_t \tilde{\mathbf{u}} + u_* \left[\frac{\theta_0}{\rho_0 u_*^2} \langle R_\alpha \rho_{\alpha*} \nabla \tilde{\rho}_\alpha \rangle + (\hat{\mathbf{u}}_0 \cdot \nabla) \tilde{\mathbf{u}} + \frac{R_0 \theta_*}{u_*^2} \nabla \tilde{\theta} \right] = \\ & = u_*^2 \left[2\tau_0 \frac{\theta_0}{\rho_0 u_*^2} (\hat{\mathbf{u}}_0 \cdot \nabla) \langle R_\alpha \rho_{\alpha*} \nabla \tilde{\rho}_\alpha \rangle + \frac{\mu_0}{\rho_0 u_*^2} \Delta \tilde{\mathbf{u}} + \right. \\ & \left. + \left(\frac{\chi_0}{\rho_0 u_*^2} + \tau_0 \frac{(R\gamma)_0 \theta_0}{u_*^2} \right) \nabla \text{div} \tilde{\mathbf{u}} + \right. \\ & \left. + \tau_0 (\hat{\mathbf{u}}_0 \cdot \nabla)^2 \tilde{\mathbf{u}} + 2\tau_0 \frac{R_0 \theta_*}{u_*^2} (\hat{\mathbf{u}}_0 \cdot \nabla) \nabla \tilde{\theta} \right], \\ & \partial_t \tilde{\theta} + u_* \left(\frac{R_0 \theta_0}{c_{V0}} \text{div} \tilde{\mathbf{u}} + \hat{\mathbf{u}}_0 \cdot \nabla \tilde{\theta} \right) = \\ & = \tau_0 u_*^2 \left[\frac{R_0 \theta_0^2}{\rho_0 c_{V0} \theta_* u_*^2} R_\beta \rho_{\beta*} \Delta \tilde{\rho}_\beta + \frac{R_0 \theta_0}{c_{V0} \theta_*} (\hat{\mathbf{u}}_0 \cdot \nabla) \text{div} \tilde{\mathbf{u}} + \right. \\ & \left. + \left(\frac{\tilde{\alpha}_0}{\rho_0 c_{V0} u_*^2} + \frac{R_0^2 \theta_0}{c_{V0} u_*^2} \right) \Delta \tilde{\theta} \right]. \end{aligned}$$

Здесь $\tau_0, \mu_0, \chi_0, \alpha_0$ – значения τ, μ, χ, α на фоновом решении. В полученных уравнениях возник оператор $(\hat{\mathbf{u}}_0 \cdot \nabla)^2 = (\hat{\mathbf{u}}_0 \cdot \nabla)(\hat{\mathbf{u}}_0 \cdot \nabla) = \hat{u}_{0i} \hat{u}_{0j} \partial_i \partial_j$, в котором предполагается суммирование от 1 до n по индексам i, j .

В полученной системе при выполнении условий

$$\frac{u_*}{\rho_{\alpha*}^2} = \frac{R_\alpha \theta_0}{\rho_{\alpha 0} \rho_0}, \alpha = \overline{1, K}, \frac{\theta_*^2}{u_*^2} = \frac{\theta_0}{c_{V0}},$$

возможна одновременная симметризация конвективных и диссипативных слагаемых. Полученные условия ниже предполагаются выполненными. Саму симметризованную систему здесь опускаем.

Исследование свойств линейаризованной системы

Введем скалярные произведения в пространствах Лебега ($L^2(\Omega)$) функций и вектор-функций в области $\Omega \subset \mathbb{R}^n$. Пусть $\mathbf{H}_0^1(\Omega)$ – замыкание в пространстве Соболева $\mathbf{H}^1(\Omega)$ гладких финитных в Ω вектор-функций. Рассмотрим полученную ранее симметризованную линейаризованную систему уравнений в цилиндре $Q := \Omega \times (0, \infty)$. Начальные и краевые условия имеют вид:

$$\tilde{\mathbf{z}}|_{\partial\Omega \times (0, \infty)} = 0, \quad \tilde{\mathbf{z}}|_{t=0} = \tilde{\mathbf{z}}^{(0)}(x).$$

Указанной системе при $\partial_t \tilde{\mathbf{z}}(\cdot, t), \nabla \tilde{\mathbf{z}}(\cdot, t) \in L^2(\Omega)$ отвечает интегральное тождество

$$(\partial_t \tilde{\mathbf{z}}, \mathbf{z})_\Omega + u_* \mathcal{B}_\Omega(\tilde{\mathbf{z}}, \mathbf{z}) + u_*^2 \mathcal{A}_\Omega(\tilde{\mathbf{z}}, \mathbf{z}) = 0$$

$\forall \mathbf{z} \in \mathbf{H}_0^1(\Omega), t > 0$. В нем для билинейных форм $\mathcal{B}_\Omega, \mathcal{A}_\Omega$ выполняются свойства

$$\mathcal{B}_\Omega(\mathbf{z}, \mathbf{z}) = 0 \quad \forall \mathbf{z} \in \mathbf{H}_0^1(\Omega), \\ \mathcal{A}_\Omega(\tilde{\mathbf{z}}, \mathbf{z}) = \mathcal{A}_\Omega(\mathbf{z}, \tilde{\mathbf{z}}) \quad \forall \tilde{\mathbf{z}}, \mathbf{z} \in \mathbf{H}^1(\Omega).$$

Положительная определенность $\mathcal{A}_\Omega(\mathbf{z}, \mathbf{z})$ обоснована в [7].

Введем пространство $\mathbf{V}(Q_T) = \{\tilde{\mathbf{z}} \in L^2((0, T); \mathbf{H}_0^1(\Omega)); \partial_t \tilde{\mathbf{z}} \in (0, T); \mathbf{H}^{-1}(\Omega)\}$, где $Q_T = \Omega \times (0, \infty)$, область Ω ограничена и $\mathbf{H}^{-1}(\Omega) = (\mathbf{H}^1(\Omega))^*$. Для начально-краевой задачи в Q введем слабое решение $\tilde{\mathbf{z}} \in \mathbf{V}(Q_T)$ для всех $T > 0$. Данное решение удовлетворяет интегральному тождеству

$$\begin{aligned} & \int_0^T \langle \partial_t \tilde{\mathbf{z}}(\cdot, t), \mathbf{z}(\cdot, t) \rangle_{\mathbf{H}^{-1}(\Omega) \times \mathbf{H}_0^1(\Omega)} dt + u_* \mathcal{B}_{Q_T}(\tilde{\mathbf{z}}, \mathbf{z}) + \\ & + u_*^2 \mathcal{A}_{Q_T}(\tilde{\mathbf{z}}, \mathbf{z}) = 0 \quad \forall \mathbf{z} \in L^2((0, T); \mathbf{H}_0^1(\Omega)) \end{aligned}$$

для всех $T > 0$.

Теорема 2. 1. Слабое решение $\tilde{\mathbf{z}} \in \mathbf{V}(Q_T)$ с любым $T > 0$ начально-краевой задачи для линейаризованной системы существует и единственно. Для данного решения верно энергетическое равенство

$$\frac{1}{2} \|\tilde{\mathbf{z}}(\cdot, t)\|_\Omega^2 + u_* \mathcal{A}_{Q_T}(\tilde{\mathbf{z}}, \tilde{\mathbf{z}}) = \frac{1}{2} \|\tilde{\mathbf{z}}^{(0)}\|_\Omega^2 \quad \forall T > 0.$$

2. Существует производная $\partial_t (\|\tilde{\mathbf{z}}\|_\Omega^2) \in L^1(0, \infty)$ и верно свойство $L^2(\Omega)$ – диссипативности

$$\partial_t (\|\tilde{\mathbf{z}}(\cdot, t)\|_\Omega^2) \leq 0 \text{ п.в. на } (0, \infty).$$

Анализ параболичности КГД системы

Для выполнения анализа параболичности системы (1)-(3) в уравнениях редуцированной системы необходимо отбросить конвективные слагаемые слева и остаточные члены $O(|\nabla \mathbf{z}|^2)$. Коэффициенты при $\partial_i \partial_j$ в рассматриваемой системе «замораживаются», к результату применяется преобразование Фурье \mathcal{F}

по x . В результате данного преобразования получаем систему

$$\partial_t \mathcal{F}z(\zeta, t) + u_*^2 |\zeta|^2 A(z_0, \xi) \mathcal{F}z(\zeta, t) = 0, t > 0, \quad (4)$$

где $\zeta \in \mathbb{R}^n \setminus \{0\}$ – параметр, а $A(z_0, \xi)$ – вещественная матрица порядка $K + n + 1$ с вектор-столбцом $\xi = \zeta/|\zeta|$. Система (4) представляет собой систему обыкновенных дифференциальных уравнений. Обозначим собственные значения введенной матрицы через $\lambda[A(z_0, \xi)]$. Далее выполним замену $z = D_* \tilde{z}$. Учитывая, что $\mathcal{F}z = \mathcal{F}D_* \tilde{z}$ и домножая (4) слева на D_*^{-1} получим эквивалентную систему с матрицей $\hat{A}(z_0, \xi) = D_*^{-1} A(z_0, \xi) D_*$, подобной матрице A . Эта матрица положительно определена.

Сформулируем теорему для системы (1)–(3) о локальной по времени однозначной разрешимости задачи Коши в слое $\Pi_T := \mathbb{R}^n \times (0, T)$ и начальных условиях $(\rho, u, \theta)|_{t=0} = (\rho^{(0)}(x), u^{(0)}(x), \theta^{(0)}(x))$ для $x \in \mathbb{R}^n$.

Пусть $\mathcal{D} := (\underline{\rho}_1, \bar{\rho}_1) \times \dots \times (\underline{\rho}_K, \bar{\rho}_K) \times (-\bar{u}_1, \bar{u}_1) \times \dots \times (-\bar{u}_n, \bar{u}_n) \times (\underline{\theta}, \bar{\theta}) \subset (0, \infty)^K \times \mathbb{R}^n \times (0, \infty)$. При $\tau, \mu, \lambda, \alpha \in C^2(\mathcal{D})$ эквивалентная КГД система, с матрицей $\hat{A}(z_0, \xi)$, удовлетворяет условию равномерной параболичности по Петровскому в \mathcal{D} [8]. Пусть $0 < \beta < 1$ – параметр.

Теорема 3. Пусть $\tau, \mu, \lambda, \alpha \in C^2(\mathcal{D})$. Пусть начальные данные $\rho^{(0)}, u^{(0)}, \theta^{(0)} \in C^{(2, \beta)}(\mathbb{R}^n)$, причем их значения $(\rho^{(0)}, u^{(0)}, \theta^{(0)})(x)$ принадлежат какому-либо компакт в \mathcal{D} и $f \in C^{(0, \beta, 0)}(\bar{\Pi}_T)$, $Q_\alpha \in C^{(1, \beta, 0)}(\bar{\Pi}_T)$, $\alpha = \overline{1, K}$. Тогда в слое Π_T и при достаточно малом $T > 0$ задача Коши для КГД системы уравнений (1)–(3) имеет единственное классическое решение $\rho, u, \theta \in C^{(2, \beta, \beta/2)}(\bar{\Pi}_T)$ с $\partial_t \rho, \partial_t u, \partial_t \theta \in C^{(0, \beta, 0)}(\bar{\Pi}_T)$, и его значения $(\rho, u, \theta)(x, t)$ принадлежат \mathcal{D} .

Здесь $C^{(m, \beta, \beta_t)}(\bar{\Pi}_T)$ с $m = 0, 1, 2, 0 \leq \beta_t < 1$ – пространство функций, удовлетворяющих условию Гёльдера порядка β по x и β_t по t (при $0 < \beta_t < 1$) равномерно в $\bar{\Pi}_T$ и имеющих непрерывные и ограниченные в $\bar{\Pi}_T$ производные порядка $k = 0, \dots, m$ по x , а $C^{(2, \beta)}(\mathbb{R}^n)$ – стандартные пространства Гёльдера.

Заключение

В работе получено уравнение баланса энтропии с неотрицательным производством энтропии для агрегированной КГД системы уравнений односкоростной и однотемпературной многокомпонентной смеси сжимаемых газов при наличии потоков диффузии. Обоснованы локальная по времени классическая однозначная разрешимость задачи Коши и параболичность по Петровскому для этой системы. Для линеаризованной системы уравнений обоснованы существование, L^2 – диссипативность и единственность слабых решений начально-краевой задачи.

Результаты работы получены совместно с проф. А.А. Злотником и опубликованы в [7]. Работа выполнена при финансовой поддержке РФФ, проект 19-11-00169.

Список литературы

1. Елизарова Т.Г. Квазигазодинамические уравнения и методы расчета вязких течений. – М.: Научный мир, 2007.
2. Пилюгин Н.Н., Тирский Г.А. Динамика ионизированного излучающего газа. – М.: Изд-во. Московского ун-та, 1989.
3. Giovangigli V. Multicomponent flow modeling. Boston, Birkhäuser, 1999.
4. Елизарова Т.Г., Злотник А.А., Четверушкин Б.Н. О квазигазо- и гидродинамических уравнениях бинарных смесей газов // Докл. АН. – 2014. – Т. 459. – № 4. – с. 395-399.
5. Елизарова Т.Г., Злотник А.А., Шильников Е.В. Регуляризованные уравнения для численного моделирования течений гомогенных бинарных смесей вязких сжимаемых газов // Ж. вычисл. матем. и матем. физ. – 2019. – Т. 59. – №. 11. – с. 1899-1914.
6. Злотник А.А., Федченко А.С. О свойствах квазигазодинамической системы уравнений гомогенной газовой смеси с общей регуляризующей скоростью // Дифференц. уравн. 2022. – Т. 58. №3. – с. 346-360.
7. Злотник А.А., Федченко А.С. Свойства агрегированной квазигазодинамической системы уравнений гомогенной газовой смеси // Докл. РАН. Матем., информ., процессы управл. – 2021. – Т. 501. – №. 1. – с. 31-37.
8. Эйдельман С.Д. Параболические системы. М.: Наука, 1964.

О ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ УГЛОВ В СЕЧЕНИЯХ МНОГОГРАННИКА СЛУЧАЙНЫМИ ПЛОСКОСТЯМИ

Константинова Е.П.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ*

Аннотация

Рассматриваются сечения многогранников случайными плоскостями. Предложен метод генерации случайных сечений и исследованы его статистические свойства. Реализован численный алгоритм оценки плотности распределения углов для выпуклых многогранников. Приведены примеры плотностей распределений в случайных сечениях для куба, правильной треугольной призмы и усеченной призмы. Разработанный алгоритм и полученные плотности распределений имеют практическое применение при анализе микроструктур кристаллических сплавов.

Введение

Твёрдые сплавы на основе карбида вольфрама широко используются в горнодобывающей промышленности. Микроснимки шлифов сплавов являются одним из основных предметов для изучения механических и физических свойств сплавов. Одним из часто применяемых методов исследования сплавов

является анализ отшлифованных сечений сплава [1-5]. В качестве примера рассмотрим твердый сплав, состоящий из зерен карбида вольфрама (WC) и кобальта (Co). Равновесной формой зёрен WC является усеченная призма (см. рис. 1). Это подтверждается как теоретически [6], так и экспериментально [7, 8]. Связующее вещество (Co) может рассматриваться как жидкий аморфный клей при спекании сплава.

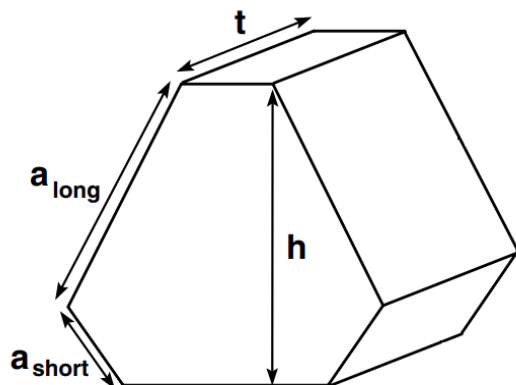


Рис. 1. Модель зерна WC

Помимо сбора традиционных параметров, таких как объёмная доля Co [9], средний размер зёрен WC [2, 4, 5], доля площади поверхности раздела фазы (contiguity) [1, 10], и средний свободный путь (mean free path) [11], сканирующие электронные микрофотографии (SEM) WC-Co можно использовать для извлечения геометрических свойств сечений, таких как углы, размер и форма регионов связующего кобальта. Компьютерный анализ микроснимков выявил интересный факт - хорошо выраженную мультимодальную плотность распределения углов между связующим веществом и зёрнами [12]. Один из пиков может быть связан с кристаллической структурой WC, а другой - с контактами зёрен. Сформулирована задача оценки плотности распределения углов, образованных сечением зерна WC случайными плоскостями. Некоторое сечение ансамбля зерен эквивалентно сечению одиночного зерна ансамблем случайных плоскостей при отсутствии контактов зёрен между собой. В работе рассчитана плотность вероятности для ансамбля сечений одного зерна случайными плоскостями. Насколько нам известно, опубликованных алгоритмов для этой задачи нет.

Алгоритм состоит из двух основных шагов - генерации плоскости со случайной ориентацией в пространстве и генерации положения секущей плоскости в многограннике. Подсчитаны углы в сечениях для оценки плотности вероятности углов. Обнаружено, что пики полученных распределений явно связаны с кристаллической формой многогранника. В качестве примеров приведены плотности распределений углов, соответствующих кубу, правильной треугольной призме, шестиугольной призме и усеченной призме. Последняя является примером зерна WC.

Генерация случайных плоскостей

Случайным вектором будем считать вектор, начало которого находится в центре сферы, а конец равномерно заметает поверхность сферы. Метод случайной генерации плоскости в пространстве будем называть *равномерным в пространстве*, если вектор нормали к сгенерированной плоскости является случайным вектором. Ориентация плоскости однозначно определяется вектором нормали, поэтому это определение приводит к равномерности ориентации плоскости.

Опишем *равномерный в пространстве* способ генерации плоскости. Необходимо генерировать вектор нормали как случайный вектор. Для того, чтобы сгенерированная точка была распределена равномерно на сфере, нужно сгенерировать два параметра вектора нормали - долготу $\theta \sim U[0, 2\pi]$ и широту $\varphi = \arccos(2l - 1)$, $l \sim U[0, 1]$. θ и l берутся из равномерного распределения U . Тогда распределение вектора с долготой и широтой θ и φ равномерно на поверхности сферы (см. [13]).

Равномерное сечение куба

Рассмотрим сечение единичного куба. Будем называть ε -кубом куб со стороной ε , полностью содержащийся внутри единичного куба. Далее рассматриваются только ε -кубы, ребра которых параллельны ребрам единичного куба. Обозначим ε -куб с центром в точке a через $K(a, \varepsilon)$. Сгенерировав k случайных плоскостей, пересекающих единичный куб, обозначим количество плоскостей, проходящих через $K(a, \varepsilon)$, через x_ε^a . Метод случайной генерации плоскости в кубе будем называть *равномерным*, ли $\forall a, b : K(a, \varepsilon), K(b, \varepsilon) \subset K : Ex_\varepsilon^a = Ex_\varepsilon^b = \text{const}$ и он *равномерный в пространстве*. Первое требование означает, что в среднем одно и то же число плоскостей проходит через любые два ε -куба.

Плоскость однозначно задаётся вектором нормали и точкой в пространстве. Выше описано, как сгенерировать вектор нормали так, чтобы способ был *равномерным в пространстве*. Осталось сгенерировать сдвиг вдоль нормали к этой плоскости. Выберем точку равномерно на отрезке AB (см. рис. 2), где отрезок AB — это проекция всех точек единичного куба на сгенерированный вектор нормали.

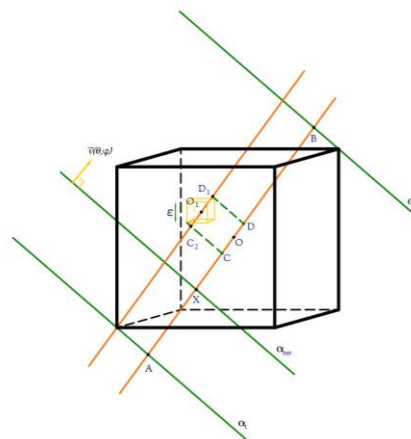


Рис. 2. Проекция всех точек куба на сгенерированную нормаль

Таким образом, любая плоскость с таким вектором нормали, которая пересекает куб, пересекает АВ и наоборот, любая плоскость с такой нормалью, которая пересекает АВ, пересекает куб. Остается доказать, что такой способ генерации является *равномерным*. Вероятность $p_a^e(\varphi, \theta)$ сгенерированной таким образом плоскости, пересекающей ε -куб, есть $|CD|/|AB|$ (см. рис. 2), где отрезок CD - это проекция всех точек ε -куба на сгенерированный вектор нормали. Покажем, что эта вероятность равна ε . Построим линию, параллельную сгенерированной, проходящую через центр ε -куба. Если обозначить через C1 и D1 концы отрезка-проекции всех точек ε -куба на новую прямую, то C1D1 = CD (см. рис. 3). Кубы и прямые подобны с отношением ε , поэтому проекции на соответствующие прямые пропорциональны с тем же отношением. $p_a^e(\varphi, \theta) = \varepsilon$ не зависит от φ , θ и а, следовательно, вероятность таким образом сгенерированной плоскости пересечь любой ε -куб, есть ε .

Так как вероятность пересечения ε -куба случайной плоскостью фиксирована, то, сгенерировав таким образом k плоскостей, оказываемся в рамках схемы Бернулли: $x_\varepsilon \sim \text{Bin}(k, \varepsilon)$. Разумеется, условие *равномерности* выполняется: $\forall a, b : K(a, \varepsilon), K(b, \varepsilon) \subset K : Ex_\varepsilon^a = Ex_\varepsilon^b = k\varepsilon$

Равномерное сечение выпуклого многогранника

Если провести аналогичное доказательство для произвольного выпуклого многогранника (см. рис. 3), то описанный метод генерации плоскости оказывается *равномерным в произвольном выпуклом многограннике*.

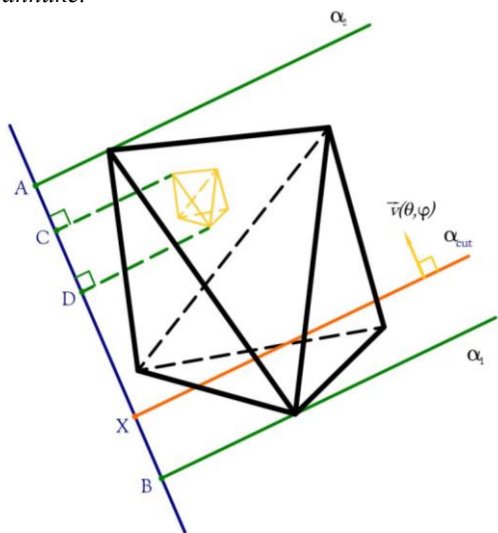


Рис. 3. Произвольный выпуклый многогранник

Оценки плотностей распределений для куба, треугольной и шестиугольной правильной призмы

Описанным способом генерируем 10^6 плоскостей сечения куба, правильной треугольной призмы с равными ребрами и правильной шестиугольной призмы с равными сторонами. Таким образом, мы получаем оценки для плотностей вероятностей углов секущих многоугольников (см. рис. 4). Для куба пик 90° , для правильной треугольной призмы пики 60° и

90° , для правильной шестиугольной призмы пики 90° и 120° .

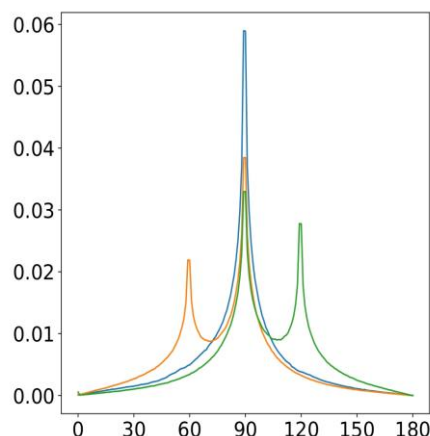


Рис. 4. Плотности распределений углов секущих многоугольников. Голубой - для куба, оранжевый - для правильной треугольной призмы, зелёный - для правильной шестиугольной призмы.

Оценки плотностей распределений для усеченных призм

Рассмотрим различные усеченные призмы, в основании которых лежит усеченный равносторонний треугольник со стороной 1 (см. рис. 1). Такая призма однозначно определяется двумя параметрами $r = a_{\text{short}}/a_{\text{long}}$ и $k = t/h$. Параметр r изменяется от 0 до 1, $r=0$ соответствует треугольнику в основании, $r=1$ соответствует шестиугольнику. Параметр k означает высоту призмы. Сгенерировав описанным методом 10^5 плоскостей, получим оценки плотностей распределений углов для различных параметров r и k (см. рис. 5). Для разных r и k графики отличаются высотой пиков 90° и 120° .

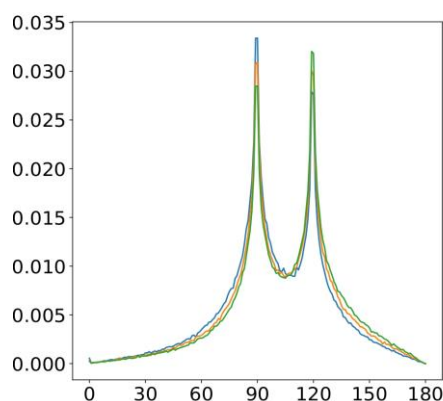


Рис. 5. Плотности распределений углов секущих усечённой призмы. Голубой - $r=1, k=0.6$, оранжевый - $r=1, k=1$, зелёный - $r=1, k=1.5$.

Заключение

Основным результатом работы является разработанный нами алгоритм, который позволяет генерировать случайные плоскости в пространстве *равномерно*. Применение этого метода позволяет получить

численную оценку плотностей вероятностей углов в случайных сечениях для куба, правильной треугольной призмы, усеченной призмы. Для сравнения, некорректный способ генерации случайных прямых в кубе, параллелепипеде и шестиугольной призме был описан для длин отрезков [14], но отсутствие равномерности в описанном выше смысле у метода генерации [14] дает сдвиг в плотностях распределений.

Список литературы

1. Gurland J. New scientific approaches to development of tool materials //International Materials Reviews. – 1988. – Т. 33. – №. 1. – С. 151-166.
2. Engqvist H., Uhrenius B. Determination of the average grain size of cemented carbides //International Journal of Refractory Metals and Hard Materials. – 2003. – Т. 21. – №. 1-2. – С. 31-35.
3. Brieseck M. et al. A straightforward method for analysing the grain-size distribution in tungsten carbide-cobalt hardmetals //Microchimica Acta. – 2010. – Т. 168. – С. 309-316.
4. Al-Thyabat S., Miles N. J. An improved estimation of size distribution from particle profile measurements //Powder Technology. – 2006. – Т. 166. – №. 3. – С. 152-160.
5. Kang K. Y., Roemer J. G., Ghosh D. Microstructural characterization of cemented carbide samples by image analysis techniques //Powder technology. – 2000. – Т. 108. – №. 2-3. – С. 130-136.
6. Christensen M. et al. Morphology of WC grains in WC-Co alloys: Theoretical determination of grain shape //Acta materialia. – 2007. – Т. 55. – №. 5. – С. 1515-1521.
7. Borgh I. et al. On the three-dimensional structure of WC grains in cemented carbides //Acta Materialia. – 2013. – Т. 61. – №. 13. – С. 4726-4733.
8. Jiménez-Piqué E. et al. Focused ion beam tomography of WC-Co cemented carbides //International Journal of Refractory Metals and Hard Materials. – 2017. – Т. 67. – С. 9-17.
9. Roebuck B., Bennett E. G. Phase size distribution in WC/Co hardmetal //Metallography. – 1986. – Т. 19. – №. 1. – С. 27-47.
10. Roebuck B. et al. Aspects of the metrology of contiguity measurements in WC based hard materials //International Journal of Refractory Metals and Hard Materials. – 2017. – Т. 62. – С. 161-169.
11. Mingard K. P. et al. Some aspects of the structure of cobalt and nickel binder phases in hardmetals //Acta Materialia. – 2011. – Т. 59. – №. 6. – С. 2277-2290.
12. Kagramanyan D. G. et al. Computer analysis of the cemented carbides' microstructure //Letters on Materials. – 2021. – Т. 11. – №. 4. – С. 447-451.
13. Muller M. E. A note on a method for generating points uniformly on n-dimensional spheres //Communications of the ACM. – 1959. – Т. 2. – №. 4. – С. 19-20.
14. Han J. H., Kim D. Y. Determination of three-dimensional grain size distribution by linear intercept measurement //Acta Materialia. – 1998. – Т. 46. – №. 6. – С. 2021-2028.

О СВОЙСТВАХ ОДНОЙ СТОХАСТИЧЕСКОЙ МОДЕЛИ ДИНАМИКИ МНЕНИЙ С ОГРАНИЧЕННЫМ РАДИУСОМ ДОВЕРИЯ

Масликов Д.Н., Шарова А.И.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В данной работе построена стохастическая модификация классической модели Хегсельмана-Краузе динамики мнений с ограниченным доверием. Доказано, что с вероятностью 1 система достигает состояния равновесия, причем за конечное время. Исследована зависимость финального профиля от параметров задачи и коэффициента возмущения. Для классической модели сформулирован алгоритм управления выделенным агентом, при котором финальный профиль всегда является консенсусом. Показано, что предложенный алгоритм управления также эффективен для стохастической модели.

Введение

Математические методы и модели для изучения процессов формирования мнений в группе агентов, проблем поляризации мнений, достижения консенсуса и т.д. начали применяться, в основном, во второй половине 20 века. В последние годы интерес к данной теме резко возрос, это связано и с задачами исследования социальных сетей, и с приложениями в биологии, робототехнике и др. Объектом нашего исследования является модель Хегсельмана и Краузе [2] (НК-модель), которая относится к классу моделей с ограниченным доверием. В их основу положен принцип гомофилии [1], а именно, схожие агенты взаимодействуют чаще, чем сильно различающиеся.

Наш интерес к НК-модели обусловлен следующим. Данная модель, с одной стороны, задается сложной нелинейной динамической системой, которая представляет интерес как самостоятельный математический объект. С другой стороны, модели с ограниченным доверием, как оказалось, имеют отношение ко многим прикладным задачам, таким, как построение алгоритмов обнаружения сообществ на графах, алгоритмов выделения кластеров в данных, байесовским алгоритмам распределенных методов принятия решений [6]. Отметим, что в НК-модели предполагается синхронность взаимодействия всех агентов, т.е., в каждый момент времени взаимодействуют и обновляют свои мнения все агенты. Такое допущение, как отмечается, реализуемо в малых группах, а в больших сообществах мало реалистично. Мы ослабляем предположения НК-модели. А именно, мы полагаем, что в каждый момент времени взаимодействует случайное число активных агентов. Активен агент или нет, определяется набором из независимых бернуллиевских случайных величин. Таким образом, предлагаемая модель является уже стохастической (StНК-модель). Мы показываем, что построенная StНК-модель сохраняет некоторые

свойства НК-модели, а также демонстрирует новые. Мы доказываем сходимость п.н. к финальному стационарному профилю за конечное время. Кроме того, мы изучаем возможности управления StHK-моделью. В данной работе мы разработали алгоритм управления, при котором в классической НК-модели финальный профиль является консенсусом, причем консенсусное значение можно сделать, как угодно, близким к 1. Отметим, что задача управления, рассматриваемая в данной работе, близка к задачам [4]–[5], однако в нашем случае класс допустимых управлений существенно уже. Разработанный алгоритм управления показывает свою эффективность и для StHK-модели.

Модель Хегсельмана-Краузе и ее свойства

Рассматривается группа из n агентов. Каждый агент характеризуется некоторым числом из отрезка $[0,1]$, число трактуется как мнение агента. Взаимодействие агентов происходит в дискретные моменты времени. Обозначим через $x_i(t)$ мнение i -го агента в момент времени t , вектор $x(t) = (x_1(t), \dots, x_n(t))$ будем называть профилем мнений в момент времени t . Зафиксируем агента i и профиль мнений $x = (x_1, \dots, x_n)$. Определим множество соседей агента i : $I(i, x) = \{1 \leq j \leq n: |x_i - x_j| \leq \varepsilon\}$, где ε - радиус доверия агента, $\varepsilon > 0$. В НК-модели агент на каждом шаге взаимодействует только с теми агентами, мнения которых попали в его область доверия, усредняя свое мнение по множеству соседей (вес каждого соседа один и тот же):

$$x_i(t+1) = |I(i, x(t))|^{-1} \sum_{j \in I(i, x(t))} x_j(t), \quad t \geq 0 \quad (1)$$

Известно [6], что отображение профилей (1) за конечное время переводит любой начальный профиль $x(0)$ в стационарную конфигурацию $x^* \in [0,1]^n$, т.е. существует такое $T < +\infty$, что $x(T+t) = x(T) = x^*$, $\forall t \geq 0$. При этом $\forall i, j$ либо $x_i^*(T) = x_j^*(T)$, либо $|x_i^*(T) - x_j^*(T)| > \varepsilon$. Множество всех таких стационарных конфигураций обозначим через M_{HK} . Среди стационарных конфигураций важную роль играют консенсусные профили x , т.е. такие, для которых все компоненты одинаковы: $x_1 = \dots = x_n$. Известны некоторые достаточные условия достижения консенсуса для НК-модели [6].

Стохастическая модель Хегсельмана-Краузе

В данной работе мы предлагаем следующую стохастическую модификацию НК-модели: пусть в каждый момент времени с некоторой вероятностью агент может стать неактивным и не участвовать в алгоритме (1) на данном шаге. Активные агенты ведут себя согласно правилам стандартной модели НК. Алгоритм взаимодействия можно описать следующим итерационным соотношением:

$$x_i(t+1) = \frac{b_{it}}{|N_i(t)|} \sum_{j \in N_i(t)} x_j(t) + (1 - b_{it})x_i(t) \quad (2)$$

где b_{it} - бернуллиевская случайная величина (с.в.), $P(b_{it} = 1) = 1 - P(b_{it} = 0) = p$. Если $b_{it} = 1$, то агент i активен в момент времени t и действует согласно правилам модели НК. Если $b_{it} = 0$, то агент i не является активным в момент времени t , не обновляет свое мнение и не влияет на мнения активных агентов. В (2) мы используем множество активных соседей агента i : $N_i(t) = \{j: b_{jt} = 1, |x_i(t) - x_j(t)| \leq \varepsilon\}$. Мы предполагаем, что с.в. в наборе $\{b_{it}\}$ независимы для всех t и всех i . В дальнейшем мы считаем, что с.в. набора $\{b_{it} = b_{it}(\omega)\}$ реализованы на некотором Колмогоровском вероятностном пространстве (Ω, \mathcal{F}, P) , причем $\omega \in \Omega$ обозначают элементарные исходы.

Таким образом, последовательность профилей $\{x(t)\}_{t=0}^{\infty}$ представляет собой марковский процесс с дискретным временем и непрерывным пространством состояний $[0,1]^n$, так как при фиксированном $x(t)$ вероятностные характеристики зависят только от параметров модели и от $x(t)$ и не зависят от прошлого процесса. На Рис. 1 приведены примеры траекторий моделей НК и StHK.

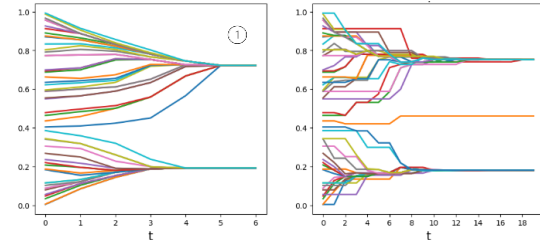


Рис. 1. Траектории НК-модели (1) и StHK-модели при $p = 0.5$ (2). $\varepsilon = 0.2$, $n = 50$

Введем с.в. $T(\omega, x(0)) = \inf\{s: x(t) \in M_{HK}\}$ - момент первого достижения случайным процессом $x(t)$, который стартует из $x(0)$, множества стационарных конфигураций M_{HK} . В работе доказано, что с.в. T конечна. Точнее, доказана теорема.

Теорема

Для всех $x(0)$ $P(\{\omega: T(\omega, x(0)) < \infty\}) = 1$.

Доказательство основано на применении условий существования предела бесконечного неоднородного произведения стохастических матриц [6], [9] и оценки сверху на время остановки в классической НК-модели [10].

В работе были проанализированы зависимости поведения StHK-модели от параметров задачи: числа агентов, радиуса доверия, вероятности активности агента. Приведем только некоторые результаты. Зафиксируем значения параметров $p = 0.5$, $n = 50$ и рассмотрим влияние радиуса доверия ε . Для каждого значения $\varepsilon = 0.1, 0.2, \dots, 1$, построим 100 реализаций для случайных начальных профилей. На Рис.2 приведены частоты образования консенсуса в финальном профиле в НК и StHK моделях. С увеличением радиуса доверия в StHK-модели, как и в НК-модели,

частота образования консенсуса q возрастает. Заметим, что зависимость q от радиуса доверия имеет пороговый характер, соответствующий гипотезе, сформулированной в [11]. На Рис. 3 приведены зависимости среднего числа кластеров в финальном профиле в НК и StНК моделях. При больших значениях ε среднее число кластеров в обеих моделях одинаково (около 1), а при малых значениях ε в StНК-модели среднее число кластеров больше. То есть, случайные возмущения вызывают усиление (в статистическом смысле) кластеризации мнений.

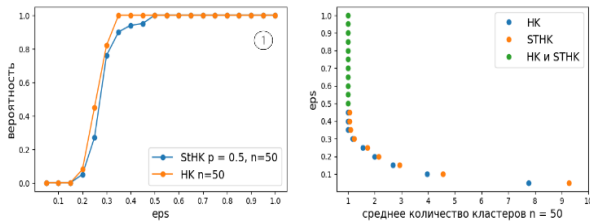


Рис. 2. Зависимость частоты консенсуса (1) и среднего числа кластеров (2) для НК и StНК моделей от ε , $n=50$, $p=0.5$

Задача управления для НК-модели

Добавим к исходной группе из n обычных агентов одного управляемого (целевого) агента (или внешнее мнение). Алгоритм взаимодействия можно описать следующим итерационным соотношением:

$$x_i(t+1) = \begin{cases} \frac{\sum_{j \in I_i(t)} x_j(t)}{|I_i(t)|}, & i \neq l \\ x_i(t) + v(t), & i = l \end{cases}$$

здесь $v(t)$ - скорость изменения мнения целевого агента в момент времени t , l - индекс целевого агента. Управляемые параметры - начальная позиция и скорость целевого агента: $x_l(0) \in [0,1]$ и $v(t) \geq 0$.

Задача управления состоит в том, чтобы максимизировать вероятность образования консенсуса и, кроме того, максимизировать консенсусное значение. Результаты моделирования показали, что эффективная начальная позиция целевого агента есть $x_l(0) = x_1 + \varepsilon$ (мы предполагаем, что $x_i(0) \leq x_{i+1}(0) \forall i$), и эффективная скорость $v(t) = \varepsilon / (1 + |I_l^-(t)|)$, где $I_l^-(t) = \{i : 1 \leq i \leq n, 0 \leq x_l - x_i \leq \varepsilon\}$. При таких значениях управляющих параметров мы гарантируем образование консенсуса и, кроме того, асимптотическое приведение профиля к 1 (Рис. 4).

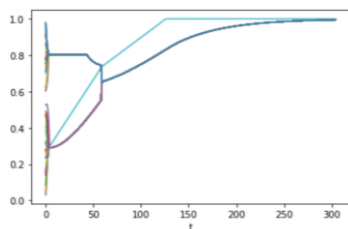


Рис. 4. НК-модель с управлением, $n = 50$, $\varepsilon = 0.2$

В работе проанализирована эффективность алгоритма управления. Для этого сравнили значения q в

НК-модели и в модели с управлением для различных значений радиуса доверия. Результаты моделирования (для 10000 случайных начальных профилей) отражены в таблице 1. Как видим, в модели с управлением консенсус достигается всегда, в отличие от НК-модели без управления.

Таблица 1. Частоты образования консенсуса q (%) в НК-модели и управляемой НК-модели, $n = 50$.

ε	0.15	0.2	0.25	0.3	0.35	0.4
НК	0.1	8.4	54.1	87.2	97.6	98.2
Control	100	100	100	100	100	100

Управление для StНК-модели

Применим к StНК-модели алгоритм управления для детерминированной НК-модели. Начальное положение целевого агента мы выбираем, как и для НК-модели: $x_l(0) = x_1(0) + \varepsilon$, а скорость целевого агента полагаем равной $v(t) = p\varepsilon / (1 + |I_l^-(t)|)$. Заметим, что скорость целевого агента уменьшается по сравнению с НК-моделью. Это связано с тем, что на каждом шаге в среднем $n(1-p)$ агентов не активны и не меняют свои мнения, то есть, появляется возможность того, что агенты выйдут из области влияния целевого агента. На Рис. 5 представлена одна из реализаций управляемой траектории StНК-модели. Результаты моделирования (Табл.2) демонстрируют эффективность предложенного алгоритма управления и для StНК-модели.

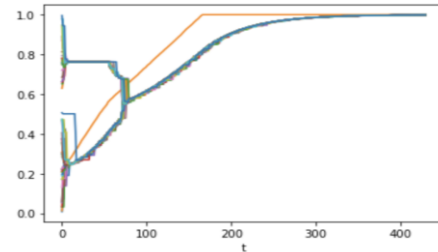


Рис. 5. Управляемая StHK-модель, $n = 50$, $p = 0.5$

Таблица 2. Значения q (%) в StНК-модели с управлением и без управления, $n = 50$, $p = 0.5$.

ε	0.15	0.2	0.25	0.3	0.35	0.4
St НК	0.01	3.1	36.2	78.2	90.6	93.5
Control	91.8	94.6	94.1	93.7	95.8	96.8

Заключение

Формирование и эволюция мнения человека – сложный процесс, зависящий от множества факторов, которые можно учитывать в модели как случайные возмущения [3], [7]-[8]. В классической НК-

модели случайность присутствует только при выборе начального профиля. Сама динамика является детерминированной. В данной работе построена стохастическая модификация НК-модели, в которой в каждый момент времени во взаимодействии участвует случайное число агентов, а именно, только те агенты, которые активны в данный момент. Для каждого агента это происходит независимо от остальных с вероятностью p . Таким образом, мы ослабляем требование синхронного взаимодействия всех участников группы. Мы показываем, что при этом сохраняется важное свойство динамики НК: конечность времени остановки, а именно, для StHK мы доказали сходимости п.н. к стационарной конфигурации за конечное время. Кроме того, мы рассмотрели задачу управления, цель которой максимизировать вероятность достижения консенсуса и обеспечить максимально возможное значение консенсуса. Предложен алгоритм, который гарантированно решает задачу для НК-модели и эффективно работает для стохастической модели.

Список литературы

1. McPherson M., Smith-Lovin L., Cook J. M. Birds of a feather: Homophily in social networks //Annual review of sociology. – 2001. – V. 27. – №. 1. – P. 415-444.
2. Hegselmann R., Krause U. Opinion dynamics and bounded confidence: models, analysis and simulation//JASSS. – 2002.-V.5.- №. 1. – P.1-33.
3. Weisbuch G., Deffuant G., Amblard F. Persuasion dynamics //Physica A: Statistical Mechanics and its Applications. – 2005. – V. 353. – P. 555-575.
4. Pineda M., Buendía G. M. Mass media and heterogeneous bounds of confidence in continuous opinion dynamics //Physica A: Statistical Mechanics and its Applications. – 2015. – V. 420. – P. 73-84.
5. Yixuan D., Cheng T., Shing W. W. Discrete-time HK-model for a leader-follower social network//2018 37th Chinese Control Conference. – IEEE, 2018. – P. 9692-9697.
6. Proskurnikov A. V., Tempo R. A tutorial on modeling and analysis of dynamic social networks//Annual Reviews in Control. – 2018. – V. 45. – P. 166-190.
7. Fotakis D., Palyvos-Giannas D., Skoulakis S. Opinion Dynamics with Local Interactions //IJCAI. – 2016. – P. 279-285.
8. Malarz K., Gronek P., Kulakowski K. Zaller-Deffuant Model of Maß Opinion//JASSS. – 2011. – V. 14. – №. 1.
9. Blondel V. D. et al. Convergence in multiagent coordination, consensus, and flocking //Proceedings of the 44th IEEE Conf. on Decision and Control. – IEEE, 2005. – P. 2996-3000.
10. Wedin E., Hegarty P. A quadratic lower bound for the convergence rate in the one-dimensional HK bounded confidence dynamics//Discrete & Computational Geometry. – 2015. – V. 53. – P. 478-486.
11. Fortunato S. Universality of the Threshold for Complete Consensus for the Opinion Dynamics of Deffuant et al//International Journal of Modern Physics C. – 2004. – V. 15. – №. 09. – P. 1301-1307.

НЕЙРОСЕТЕВОЙ АНАЛИЗ ФАЗОВЫХ ПЕРЕХОДОВ В АНИЗОТРОПНОЙ МОДЕЛИ ИЗИНГА

Суховерхова Д.Д.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

Проведено исследование анизотропной модели Изинга методами машинного обучения. Исследуется возможность нейронных сетей извлекать информацию о фазовом переходе в модели Изинга, обученных на квадратной решетке, и переносить полученные знания при тестировании модели Изинга с анизотропией. Сделан вывод об обобщающей способности нейронных сетей и корректности полученных предсказаний: предложенный метод анализа фазовых переходов в модели Изинга с анизотропией справляется с качественным определением критической температуры, но полученные таким образом результаты обладают значительной погрешностью.

Введение

В работе рассматривается модель Изинга [1] на квадратной решетке размера $L \times L$ с периодическими граничными условиями без внешнего магнитного поля, с гамильтонианом, заданным следующим уравнением [2]:

$$\mathcal{H} = -J \sum_{x,y=1}^L \sigma_{x,y} (\sigma_{x+1,y} + \sigma_{x,y+1})$$

При нарушении симметрии система становится анизотропной. Мы рассматриваем две модификации исходной модели.

Первая модель получена добавлением взаимодействия между диагональными соседями, таким образом мы получаем треугольную решетку. Диагональное взаимодействие спинов обеспечивается коэффициентом J_d . Без ограничения общности положим $J = 1$ и $J_d \in (-1, 0]$.

Энергия конфигурации с анизотропией на треугольной решетке задается гамильтонианом [3, 4]:

$$\mathcal{H} = - \sum_{x,y=1}^L [J \sigma_{x,y} (\sigma_{x+1,y} + \sigma_{x,y+1}) + J_d \sigma_{x,y} \sigma_{x+1,y+1}]$$

Для анизотропной модели Изинга известно точное соотношение для температуры фазового перехода при заданных коэффициентах взаимодействия J и J_d [3]:

$$\left(\sinh \frac{2J}{k_B T_c} \right)^2 + 2 \sinh \frac{2J}{k_B T_c} \sinh \frac{2J_d}{k_B T_c} = 1$$

Вторая модель - модель Изинга на квадратной решетке с различными константами взаимодействия по горизонтали (J_h) и вертикали (J_v) [5]:

$$\mathcal{H} = - \sum_{x,y=1}^L [J_h \sigma_{x,y} \sigma_{x+1,y} + J_v \sigma_{x,y} \sigma_{x,y+1}]$$

Точное решение [5]:

$$\sinh \frac{2J_h}{k_B T_c} \sinh \frac{2J_v}{k_B T_c} = 1$$

Во всех рассматриваемых моделях фазовый переход из ферромагнитной в парамагнитную фазу характеризуется ростом корреляционной длины. Гипотеза заключается в том, что методами машинного обучения, в частности, нейронными сетями, можно извлечь эту информацию из обучающих данных и определять критическую температуру фазового перехода в моделях из того же класса универсальности. Целью данной работы является проверка этой гипотезы.

Данные

Обучающий и тестовый наборы данных для нейронной сети состоят из черно-белых "изображений" - матриц размера $L \times L$, каждый элемент которой имеет значение +1 или -1. Для генерации данных используется алгоритм Метрополиса [2].

Время термализации, согласно [6], составляет $20L^{2.15}$. После процесса термализации каждое изображение сохраняется один раз в $2L^{2.15} \text{ mcsweep}$ [7]. Таким образом конфигурации модели Изинга генерируются для каждой комбинации J_d и L . Рассматриваемые размеры решеток: $L=20,30,40,60,80,120$; значения J_d - от -0.7 до 0.0 с шагом 0.1.

Значения параметров модели с отличающимися константами взаимодействия: $J_h = 1.0$, $J_v \in [1.0, 0.75, 0.5, 0.25, 0.125, 0.0625]$, $M = \frac{J_v}{J_h}$.

Каждое изображение генерируется для своего значения температуры. Для тренировочного набора данных для каждой температуры генерируется 2048 примеров. Для тестового набора данных для каждого из температурных точек генерируется 512 изображений.

Также для каждого изображения сохраняются метки класса, полученные на основе значений критической температуры из точного решения:

- 0 - если температура конфигурации ниже критической
- 1 - если температура конфигурации выше критической

Детали обучения

Нейронная сеть состоит из следующих слоев:

- Сверточный слой Conv2d с 1 входным каналом и 64 выходными, размер свертки 2×2 , шаг (stride) свертки = 1,
- Слой MaxPool2d 2×2 ,
- Слой активации ReLU,
- Полносвязный (linear) слой с $64 \times (L/2-1) \times (L/2-1)$ нейронами на входе и 64 на выходе,

- Слой активации ReLU,
- Полносвязный (linear) слой с 64 входными нейронами.

Поскольку параметры нейронной сети зависят от размера решетки, для каждого значения L создается отдельная модель, которая обучается на изображениях одинакового размера. Поэтому далее под словами "модель" и "сеть" будет подразумеваться совокупность нейронных сетей для разных значений L .

Последний полносвязный слой имеет на выходе 1 нейрон - результатом предсказания является вероятность принадлежности классу 1, то есть является ли температура, при которой было сгенерировано входное изображение, выше критической.

Перед обучением тренировочные данные перемешиваются и разделяются на "пакеты" (батчи) - по 4 изображения на каждую итерацию обучения нейронной сети. Тестовые данные не перемешиваются, так как результаты тестирования не зависят от порядка получения тестовых данных, и данные передаются по одному изображению.

В качестве функции потерь выбрана бинарная кросс-энтропия:

$$Q(f_i, \hat{f}_i) = -\frac{1}{n} \sum_{i=1}^n [f_i \cdot \log \hat{f}_i + (1 - f_i) \cdot \log (1 - \hat{f}_i)]$$

где f_i - правильный ответ, \hat{f}_i - предсказание сети, n - количество изображений.

В качестве оптимизации весов выбран алгоритм Adam [8], $\eta_t = 10^{-4}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$.

Обучение происходит только на конфигурациях на квадратной решетке. Таким образом, при оптимизации параметров нейронная сеть не получает информацию о состояниях с нарушенной симметрией решетки; в тестовой выборке же находятся изображения, соответствующие конфигурациям с анизотропией.

Результаты

Для определения критической температуры были рассмотрены усредненные предсказания сети:

$$F^T = \frac{1}{N} \sum_{i=1}^N f_i^T$$

где $N=512$, $j=1,2,...,M$ - номер температуры из выбранного диапазона.

Также были исследованы погрешности измерений: так как в тестовом наборе данных на каждое значение температуры приходится по 512 изображений, то можно посчитать стандартное отклонение предсказаний сети для каждой температуры:

$$V^T = \sqrt{\frac{1}{N} \sum_{i=1}^N (f_i^T)^2 - \left(\frac{1}{N} \sum_{i=1}^N f_i^T \right)^2}$$

Ожидается, что нейронная сеть будет с высокой точностью определять классы состояний, сгенерированных при температурах, далеких от критической: при низких температурах предсказания сети должны принимать значение 0, а при высоких = 1. В области же критической температуры происходит переход между двумя состояниями, который характеризуется ростом корреляционной длины до бесконечности (в случае конечного размера решеток - до L). В этой области конфигурации выше и ниже критической температуры "похожи", то есть ожидаемый выход сети для таких данных = 0.5.

Для состояний без анизотропии (рис. 1) можно увидеть, что полученные результаты совпадают с ожидаемыми: вблизи критической температуры предсказания сети приблизительно равны 0.5. При этом, при температурах, находящихся на краю рассматриваемого интервала, сеть ошибается редко и верно предсказывает вероятность быть выше критической температуры.

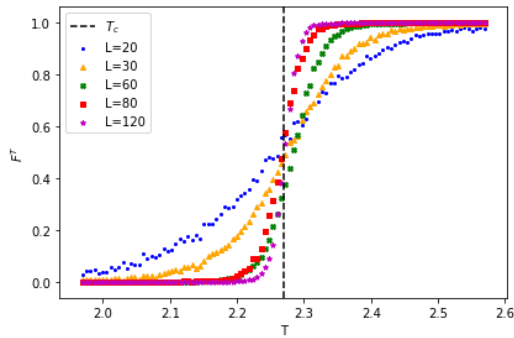


Рис. 1. Выходы сети для модели Изинга на квадратной решетке в диапазоне $[T_c - 0.3, T_c + 0.3]$, 100 температурных точек; вертикальной линией обозначена критическая температура $T_c = 2.269$

Результаты для модели Изинга на треугольной решетке при температурах выше и ниже новой критической сохраняются, однако, наблюдается следующая тенденция: средние значения ответов сети в критической температуре снижаются, и максимумы погрешностей предсказаний смещаются вправо. Также при увеличении $|J_d|$ сеть определяет критическую температуру правее ее теоретического значения.

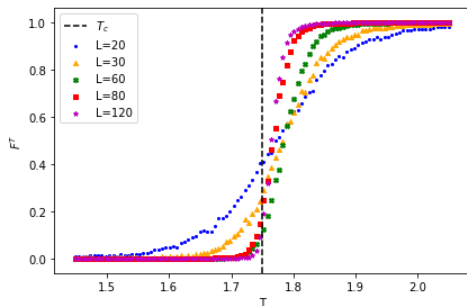


Рис. 2. Выходы сети для модели Изинга с анизотропией на треугольной решетке ($J_d = -0.3$) в диапазоне $[T_c - 0.3, T_c + 0.3]$, 100 температурных точек; вертикальной линией обозначена критическая температура $T_c = 1.749$

Определенные нейронной сетью значения температуры фазового перехода T^\wedge и их отклонения от критической температуры, найденной из точного решения, представлены на рис. 3 и рис. 4.

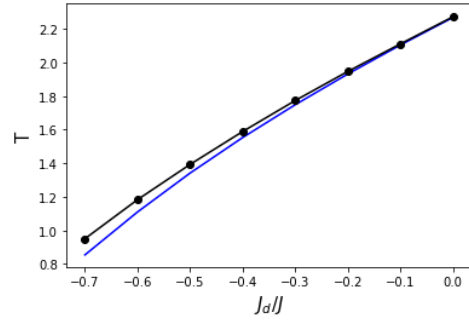


Рис. 3. Температуры фазового перехода: синей линией обозначена критическая температура, найденная из точного решения (T_c), точками обозначена критическая температура, определенная сетью (T^\wedge)

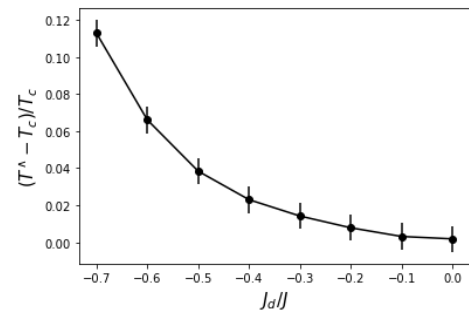


Рис. 4. Относительная погрешность определения T^\wedge

Результаты тестирования сети на данных модели Изинга с различными силами связи по вертикали и диагонали и $M = 0.125$ показаны на рис. 5.

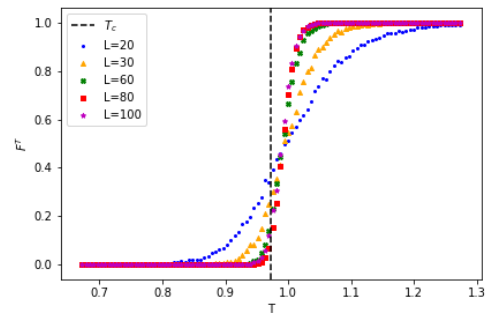


Рис. 5. Усредненные предсказания сети, $M = 0.125$ в диапазоне $T_c \pm 0.3$ для 100 температурных точек; вертикальной линией обозначена критическая температура = 0.972

С уменьшением силы взаимодействия (M) температура фазового перехода T^* также уменьшается. При этом сохраняется тенденция: определенная сетью точка фазового перехода T^* располагается правее от T_c , и разница $|T_c - T^*|$ увеличивается по мере уменьшения M (рис. 6). При этом отметим, что отно-

сительная погрешность в этом случае меньше, чем при треугольной решетке.

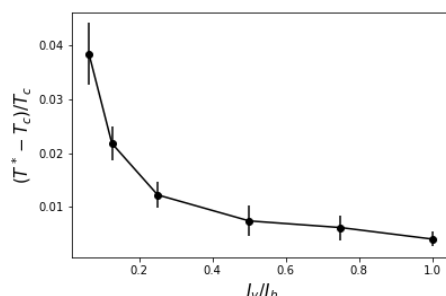


Рис. 6. Относительная погрешность определения T^*

Заключение

В работе рассмотрена модель Изинга с анизотропией. Методами машинного обучения были определены точки фазового перехода при помощи нейронной сети, обученной на данных модели Изинга без анизотропии.

Нейронная сеть, обученная на конфигурациях модели Изинга на квадратной решетке, способна качественно уловить рост корреляционной длины вблизи точки фазового перехода в незнакомых данных. При этом нарушение симметрии системы (добавление диагональных взаимодействий или изменение силы взаимодействия спинов по разным направлениям) приводит к тому, что нейронная сеть определяет точку фазового перехода с относительной погрешностью до 12%. Такая ошибка не позволяет говорить о том, что нейронная сеть не способна в точности обнаружить рост корреляционной длины в данных, которые не находились в обучающей выборке.

Основной результат работы состоит в том, что нейронная сеть предсказывает систематическое отклонение критической температуры от точной, что указывает на неуниверсальность применения метода машинного обучения для извлечения температуры фазового перехода. В том же время, анизотропия не влияет на точность оценки критического индекса корреляционной длины, что указывает на применимость метода машинного обучения для извлечения информации о критических индексах и класса универсальности модели [9].

Список литературы

1. Ising, E. Beitrag zur Theorie des Ferromagnetismus // *Z. Phys.*, vol. 31, pp. 253—258, Feb. 1925.
2. Janke W. Monte Carlo methods in classical statistical physics // *Computational many-particle physics*. – 2008. – С. 79-140.
3. Houtappel R. M. F. Order-disorder in hexagonal lattices // *Physica*. – 1950. – Т. 16. – №. 5. – С. 425-455.
4. Selke W., Shchur L. N. Critical Binder cumulant in a two-dimensional anisotropic Ising model with competing interactions // *Physical Review E*. – 2009. – Т. 80. – №. 4. – С. 042104.

5. Onsager L. Crystal statistics. I. A two-dimensional model with an order-disorder transition // *Physical Review*. – 1944. – Т. 65. – №. 3-4. – С. 117.

6. Sokal A. Monte Carlo methods in statistical mechanics: foundations and new algorithms. – Springer US, 1997. – С. 131-192.

7. Chertakov V., Shchur L. Universality classes and machine learning // *Journal of Physics: Conference Series*. – IOP Publishing, 2021. – Т. 1740. – №. 1. – С. 012003.

8. Kingma D. P., Ba J. Adam: A method for stochastic optimization // *arXiv preprint arXiv:1412.6980*. – 2014.

9. Sukhoverkhova, D., Chertakov, V., Burovski, E., Shchur, L. Validity and limitations of the supervised learning for phase transition research // *RuSCDays, 2023* (unpublished).

МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННО-РАСПРЕДЕЛЕННЫХ ЭВОЛЮЦИОННЫХ ИГР СО СЛУЧАЙНОЙ СМЕНОЙ СТРАТЕГИЙ

Антонов Д.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В игре Новака-Мэя агенты играли в эволюционную игру с соседями и основываясь выбирали стратегию соседа с наибольшим выигрышем. Мы изучаем модификацию, в которой агент выбирает следующую стратегию с вероятностью, зависящей от разности выигрышей противоположных стратегий.

Введение

Теория игр описывает социальные дилеммы, например, ситуации, при которых каждый агент преследует свои наилучшие интересы, что приводит к ухудшению общего результата. Достаточно много сложных проблем, с которыми мы сталкиваемся, от межличностных до международных, по своей сути являются социальными дилеммами [4].

Новак и Мэй [1]-[2] исследовали влияние пространственной структуры популяции на кооперацию агентов. Они обнаружили, что опеределенные структуры приводят к кооперативному поведению, которое невозможно при игре всех-со-всеми. В этой модели агенты действуют детерминировано, рационально и выбирают ту стратегию, которая принесла наибольший выигрыш. В реальном мире агенты чаще всего ведут себя рационально. Один из способов добавить иррациональность в поведение агентов реализуется через случайность в смене стратегии. В работе 1994 года Новак, Мэй и Бенхоффер исследовали модификацию игры, в которой вероятность играть определенную стратегию была равна отноше-

нию суммарного выигрыша соседей, использовавших данную стратегию, к общему выигрышу. Жабо и др.[3] исследовали случай асинхронной игры, в которой вероятность смены стратегии зависела от разницы максимальных выигрышей стратегий среди соседей агента.

В настоящей работе мы исследуем влияние случайной смены стратегий в случае, когда вероятности смены стратегии зависят от разницы выигрышей. Это может иметь влияние на уровень кооперации в пространственной игре.

Рассмотрим игру Новака-Мэя на квадратной сетке $L \times L$ с периодичными граничными условиями с L^2 агентами. При взаимодействии друг с другом, они могут либо кооперировать (кооператоры, C), либо предавать (дефекторы, D), получая выигрыш, указанный в таблице 1.

Таблица 1. Выигрыши агентов при взаимодействии

Агент 1	Агент 2	
	D	C
D	0, 0	b, 0
C	0, b	1, 1

В игре Новака-Мэя[1]-[2], каждый агент попарно взаимодействует с 8 соседями и самим собой, после чего выбирает поведение (кооперировать или предавать), принесшее наибольший суммарный выигрыш на текущем ходу среди себя и соседей и игра повторяется. Агенты не имеют памяти, поэтому принимают решение, основываясь на текущих данных.

Выигрыш агента x в результате взаимодействий в раунде t можно найти как

$$U_x(t) = \sum_{y \in n.n.} \mathbf{s}_x^T H \mathbf{s}_y, \quad (1)$$

где \mathbf{s}_x и \mathbf{s}_y - вектора стратегий агентов x и y , такие, что кооператору (C) соответствует вектор $(1, 0)^T$, а дефектору (D) вектор $(0, 1)^T$ и суммирование производится по соседям агента x .

Матрица выигрыша H имеет вид

$$H = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix}$$

Новак и Мэй обнаружили, что локальные взаимодействия в пространстве сами по себе могут обеспечить кооперативное поведение, которого не было бы в однородных популяция [1]-[2].

В данной статье мы изучаем влияние неопределенности при смене стратегии на уровень кооперации.

Игра со случайной сменой стратегии

Рассмотрим квадратную сетку $L \times L$ с периодичными граничными условиями с L^2 агентами. На каждом шаге каждый агент x взаимодействует со своими соседями и получает выигрыш по формуле (1), после чего агент меняет свою стратегию с вероятностью, зависящей от разности максимальных выигрышей соседей кооператора и дефектора. Все агенты меня-

ют стратегии одновременно, после чего игра повторяется.

Рассмотрим агента x . Пусть U_x^C и U_x^D - максимальные выигрыши кооператора и дефектора среди соседей и самого агента x

$$U_x^C = \max\{U_y: y \in n.n. \cup \{x\} \text{ и } s_y = C\}$$

$$U_x^D = \max\{U_y: y \in n.n. \cup \{x\} \text{ и } s_y = D\}$$

Если агент x в момент времени t был кооператором, то на следующем шагу он станет дефектором с вероятностью

$$W[C \rightarrow D] = 1/(1 + \exp[(U_x^C - U_x^D)/K]),$$

и наоборот, если агент был дефектором, то станет кооператором с вероятностью

$$W[D \rightarrow C] = 1/(1 + \exp[(U_x^D - U_x^C)/K]),$$

где K - параметр³.

Заметим, что при $K \rightarrow 0$ данное правило изменения стратегии совпадает с игрой Новака-Мэя, так как $W[C \rightarrow D] \rightarrow 1$ при $U_x^C < U_x^D$ и $W[C \rightarrow D] \rightarrow 0$ при $U_x^C > U_x^D$.

Похожая модель рассматривалась в работе Жабо [3] для случая асинхронного обновления стратегий.

Средняя плотность кооператоров по времени вычисляется по формуле

$$f_c^T = \frac{1}{T-\tau} \sum_{t=\tau}^T f_c(t),$$

где $f_c(t)$ - плотность кооператоров на шаге t , τ - шаг, начиная с которого считаем среднее, T - общее число шагов.

Результаты

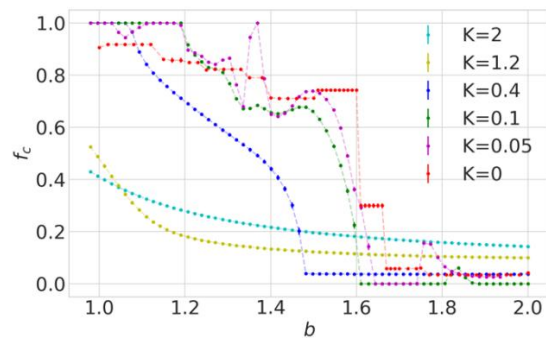


Рис. 1. Зависимость средней плотности кооператоров $\langle f_c^T \rangle$ от параметра игры b . Средняя плотность вычисляется на интервале $[\tau, T]$, с $\tau = 13500$ и $T = 15000$

На рис. 1 представлена зависимость плотности кооператоров от параметра игры при различных значениях параметра K . Добавление случайности делает зависимость средней плотности плавным, убирая резкие переходы и уменьшает долю кооператоров

при $1.2 < b < 1.65$ и, наоборот, повышает долю кооператоров при $b > 1.65$ для $K > 0.1$ и $b < 1.2$ при $K \leq 0.1$.

Так же, при $K = 0.05, 0.1$ наблюдается рост доли кооператоров при $b \approx 1.35, 1.5$ и 1.8 . Для выявления причин такого поведения требуется дальнейшее исследование.

Заключение

В данной работе было проведено исследование модификации игры Новака-Мэя со случайным изменением стратегии.

Добавление случайности при смене стратегии приводит к изменению доли кооператоров в популяции при различных значениях выигрыша и позволяет как поощрять, так и подавлять кооперацию среди агентов.

Постановка задачи и руководство работой - Е.А. Буровский.

Вычисления проводились с использованием Суперкомпьютерного комплекса НИУ ВШЭ.

Список литературы

1. Martin Nowak and Robert May, Evolutionary games and spatial chaos, Nature 359(10), 826(1992).
2. Martin Nowak and Robert May, The spatial dilemmas of evolution, International Journal of Bifurcation and Chaos 03(01), 35(1993).
3. György Szabó, Jeromos Vukov and Attila Szolnoki, Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices, Phys. Rev. E 72, 047107 (2002).
4. P. Kollack, Social dilemmas: The anatomy of cooperation, Annual Review of Sociology, vol. 24, pp. 183–214 (1998).
5. Martin Nowak, Sebastian Bonhoeffer and Robert May, More spatial games, International Journal of Bifurcation and Chaos 04(01), 33(1994).

МЕТОДИКА РАСЧЕТА ПАРЦИАЛЬНЫХ ДАВЛЕНИЙ СИСТЕМЫ: ЭТАНОЛАМИНОВОГО РАСТВОРА – H₂S-CO₂

Денисова А.И.

*Казанский национальный исследовательский технологический университет «КНИТУ-КХТИ»,
кафедра машин и аппаратов химических
производств ИХНМ*

Аннотация

Основной задачей исследования является разработка и совершенствование математического моделирования технологических процессов этаноламиновой очистки углеводородов, с целью повышения энергоэффективности производства. Предложена авторская методика расчета парциальных давлений в

системе H₂S-CO₂-H₂O-МЭА и проведен ряд математических вычислений, доказывающий адекватность разрабатываемой методики на основе сравнения с традиционным методом расчета. Разрабатываемая методика, дает возможность получения более точных результатов расчета, которые, в свою очередь играют важную роль в системном подходе усовершенствования математической модели и эксперимента. Успешное слияние этих двух составляющих, позволяют оптимизировать не только конструктивные параметры, но и весь процесс в целом.

Введение

Одна из наиболее сложных и важных задач для нефтяных перерабатывающих заводов (НПЗ), это очистка углеводородного сырья от сернистых соединений. В этой связи, разработка и усовершенствование методов расчета процессов сероочистки, а также их моделирование не теряют своей актуальности и по сей день.

Опираясь на мировой опыт, можно сказать, что наибольшее применение для сероочистки углеводородного сырья, нашли абсорбционные регенеративные процессы этаноламиновой очистки, для данных процессов применяется около -70 % установок от общего числа установок, используемых для различных методов очистки углеводородного сырья [1]. Это обусловлено наличием ряда факторов, таких как: высокая поглотительная способность этаноламинов по отношению к кислым компонентам, относительная летучесть поглотительного раствора, возможность регенерации, взаимная растворимость абсорбента с углеводородами и доступность их производства.

Можно сделать вывод, что процессы этаноламиновой очистки газов от H₂S, CO₂, меркаптанов и других серосодержащих примесей, относятся к наиболее типичным примерам применяемых в хемосорбционных процессах.

Главной особенностью в гетерогенных системах данного типа, является одновременное протекание явлений диффузионного переноса вещества из газовой фазы в объеме жидкой фазы и химического взаимодействия переносимого компонента с активной частью хемосорбента. Поэтому при строгом описании хемосорбционного процесса, необходимо учитывать кинетику химического взаимодействия реагирующих веществ и использование тем самым, кинетических моделей.

Одновременное рассмотрение физической и химической кинетики достаточно затруднено и в инженерном применении требует введение дополнительных ограничений и допущений. Особенно большие сложности возникают при описании одновременной абсорбции H₂S и CO₂.

Все это введет к тому, что технологические параметры процесса, апробированные для очистки газов от CO₂, переносятся на процесс очистки от H₂S и наоборот, а математическое описание переноса в двухфазной системе замыкается на поверхности раздела фаз.

При этом целый ряд допущений и ограничений, который в этом случае приходится вносить в математическое описание, существенно снижает точность вычислений парциальных давлений. Особенно большие сложности возникают при описании одновременной абсорбции H₂S и CO₂. Исходя из этого, на практике, чаще всего, применяемые модели, не учитывают химическое взаимодействие, опираясь при этом только на физический массоперенос, стоит отметить, что это является необходимой частью описания процесса, который строится на кинетических закономерностях.

Таким образом, при любой постановке задачи, применение описания парожидкостного равновесия оказывает принципиальное значение на результаты моделирования процесса. В этих условиях представляется целесообразным разработки методики расчета равновесного состояния системы на основе статистического анализа результатов расчета в широких интервалах изменения состояния системы (P, X, T) при различной концентрации поглотительного раствора.

Математическое описание системы

Основной задачей моделирования является расчет зависимости парциального давления CO₂ над раствором от количества растворенного углекислого газа. Количество CO₂, поглощенного раствором амина, без учета химических реакций можно рассчитать по уравнениям состояния Пенга–Робинсона (Peng–Robinson) [2] или Соаве–Редлиха–Квонга (Soave–Redlich–Kwong) [3]. Количество поглощенного CO₂ с учетом химических реакций может быть рассчитано с использованием модели Kent–Eisenberg [4] или Li–Mather [5].

Модель Kent–Eisenberg является эмпирической, основанной на экспериментальных данных о растворимости CO₂ в различных растворах аминов при различных давлениях и температурах. Как правило, полученные результаты, обрабатываются линейным регрессионным анализом [6]. Чтобы оценить полученный результат, применяется метод наименьших квадратов.

Исходя из этого, большой интерес представляет разработка математических моделей, обобщающих экспериментальные данные, с помощью которых определяются зависимости парциальных давления кислых компонентов над раствором аминов от состава раствора и температуры.

В данном исследовании, для описания равновесия с позиции закона действующих масс для электролитов, применяется уравнение:

$$P_{H_2S} = \frac{K_W}{k_{1c} \cdot k_i \cdot H} \cdot \frac{\gamma^2}{a} \cdot \frac{[H\bar{S}][RNH_3]}{[RNH_2]} \quad (1)$$

P_{H_2S} – Давление H₂S над раствором (мм.рт.ст.);

K_W – Константа равновесия диссоциации воды (моль);

K_{1c} – Константа равновесия диссоциации H₂S (моль);

K_i – Константа равновесия ионизации этаноламина (моль);

H – Константа Генри для системы H₂S–H₂O (моль/1000 г H₂O мм.рт.ст.);

δ – Мольная доля для воды в растворе;

α – Коэффициент, учитывающий активность свободного ЭА по сравнению с водой;

γ – Средний коэффициент активности ионов;

$[H\bar{S}]$, $[RN_3^+]$, $[RNH_2]$ – концентрация гидросульфид ионов, ионизированного и свободного ЭА [моль/моль].

Уравнение (1) может быть выражено через общую концентрацию сероводорода – X [моль H₂S/моль МЭА] и общую концентрацию этаноламина – m [моль ЭА/1000г H₂O]:

$$P_{H_2S} = \frac{K_W}{K_{1c} \cdot K_i \cdot H} \cdot \frac{\gamma^2}{a \cdot \delta} \cdot \frac{m \cdot X^2}{1 - X} \quad (2)$$

Надо отметить, что данные уравнения не учитывают концентрации фактически растворенного сероводорода и сульфид ионов.

Для этого были исследованы девять экспериментальных точек, значения которых взяты из источников литературы.

В качестве базового описания были выбраны уравнения (1) - (2). Однако для случая одновременной абсорбции H₂S и CO₂, под концентрацией X следует понимать сумму концентрация $X_{H_2S} + X_{CO_2}$. Было принято учитывать концентрацию ионов $[HCO_3^-]$, при насыщении раствора CO₂ свыше 0,25 моль на общее количество этаноламина.

Для интервала $0 \leq \sum X_i \leq 0.25$

$$P_{H_2S} = \frac{K_W \cdot \gamma^2 \cdot m}{K_i \cdot K_{1c} \cdot H_{H_2S} \cdot \alpha \cdot \delta} \cdot \frac{X_{H_2S} \cdot (X_{H_2S} + X_{CO_2})}{(1 - X_{H_2S} - 2X_{CO_2})} \quad (3)$$

$$P_{CO_2} = \frac{\gamma^2}{K_M \cdot a^2 \cdot \delta^2} \cdot \frac{X_{CO_2} \cdot (X_{H_2S} + X_{CO_2})}{(1 - X_{H_2S} - X_{CO_2})^2} \quad (4)$$

Для интервала $0.25 \leq \sum X_i \leq 0.7$

$$P_{H_2S} = \frac{K_W \cdot \gamma^2 \cdot m}{K_i \cdot K_{1C} \cdot H_{H_2S} \cdot \alpha \cdot \delta \cdot X_{H_2S} \cdot (X_{H_2S} \cdot X_{CO_2})} \cdot \frac{[HCO_3^-]}{(\frac{[HCO_3^-]}{m} + 1 - X_{H_2S} - 2X_{CO_2})} \quad (5)$$

$$P_{CO_2} = \frac{\gamma^2}{K_M \cdot \alpha^2 \cdot \delta^2} \cdot \frac{(X_{CO_2} - \frac{[HCO_3^-]}{m}) \cdot (X_{H_2S} + X_{CO_2})}{(\frac{[HCO_3^-]}{m} + 1 - X_{H_2S} - X_{CO_2})^2} \quad (6)$$

$$\frac{[HCO_3^-]}{m} = \sqrt{\frac{K_1}{(\frac{K_1}{\alpha \cdot m} + 1 - X_{H_2S} - 2X_{CO_2})^2} + \frac{K_i \cdot X_{CO_2}}{\delta \cdot m}} - \frac{\frac{K_i}{\alpha \cdot m} + 1 - X_{H_2S} - 2X_{CO_2}}{2} \quad (7)$$

Приведенные зависимости, результаты расчета изменения параметров состояния, которых, отображены в табл.1. показали, что отклонение расчетных значений парциальных упругостей кислых газов над растворами этаноламинов удовлетворительно согласуются с экспериментальными значениями, достигая высокой сходимости на отдельно взятых интервалах.

Исходя из этого можно судить об адекватности модели, и продолжить исследование в данном направлении.

Таблица 1. Сравнение расчетных и экспериментальных данных P_{H_2S} и P_{CO_2}

T, °C	100	100	100	100	120
X_{CO_2}	0,38	0,18	0,023	0,07	0,32
$P_{CO_2}(\text{эксп.})$	217	15	30	77	328
$P_{CO_2}(\text{расч.})$	124,5	10,1	8,69	91,8	376,4
X_{H_2S}	0,005	0,007	0,04	0,548	0,005
$P_{H_2S}(\text{эксп.})$	8,6	2,1	17	932,4	6,6
$P_{H_2S}(\text{расч.})$	8,8	2,1	20,5	870	6,2

Таблица 1.1. Продолжение

T, °C	60	60	40	40
X_{CO_2}	0,0016	0,038	0,029	0,047
$P_{CO_2}(\text{эксп.})$	0,8	14,2	7,8	13,3
$P_{CO_2}(\text{расч.})$	0,34	1,39	1,54	1,91
X_{H_2S}	0,0107	0,088	0,35	0,027
$P_{H_2S}(\text{эксп.})$	5,9	23,2	76,4	5,4
$P_{H_2S}(\text{расч.})$	1,67	11,8	27	6,1

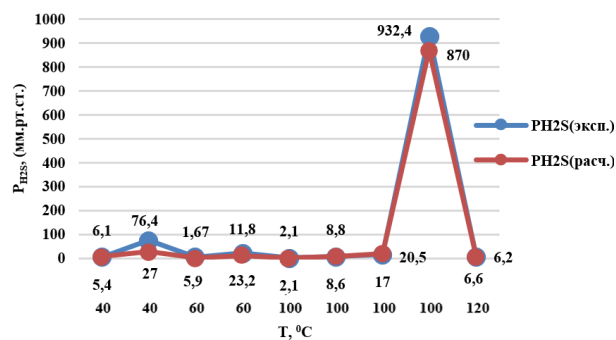


Рис. 1. Сравнение расчетных и экспериментальных значений PCO_2

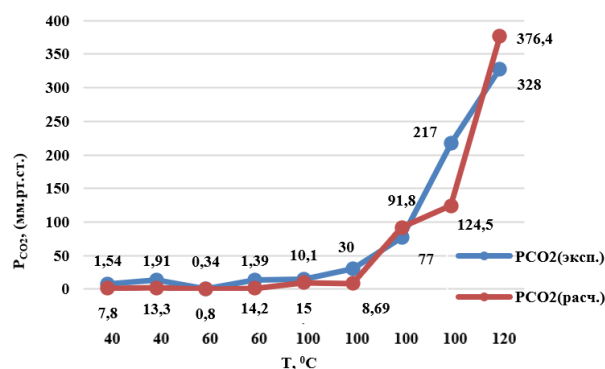


Рис. 2. Сравнение расчетных и экспериментальных значений PH_2S

Данная модель фазового равновесия позволяет проводить прямой расчет парциальных давлений H_2S и CO_2 над растворами этаноламинов при заданных параметрах состояния. Модель позволяет выполнять решение и обратной задачи определение концентрации при заданных значениях P_{CO_2} и P_{H_2S} методом последовательных приближений.

На данном этапе весь расчет парциальных давлений системы этаноламинового раствора H_2S - CO_2 был произведен с помощью программы – Microsoft EXCEL с поддержкой макроса, а также в MATHCAD.

Модель и алгоритм расчета упругости насыщенных паров

Как отмечалось ранее, использование допущений при описании соотношения между модулем сульфидирования и карбонизации раствора, является одним из недостатков расчета.

$$\beta = \frac{X_{H_2S}}{X_{CO_2}} \quad (8)$$

Данное соотношение переменного по высоте аппарата, определяется всем комплексом взаимосвязан-

ных параметров технологического процесса и не может быть задано априорно.

Главное влияние на формирование соотношения в процессе хемосорбции (десорбции) будут оказывать кинетические закономерности химического взаимодействия кислых компонентов с этанолами́нами. В частности, есть рекомендации принимать β равным соотношению между концентрациями H_2S и CO_2 в исходном сырье.

Следует отметить еще один аспект использования методики расчета парциальных упругостей, связанный с необходимостью задания степеней сульфидирования и карбонизации растворов.

В реальных расчетах разделения чаще приходится решать задачу в обратной постановке: по заданному парциальному давлению кислых компонентов в газовой фазе определять равновесную степень карбонизации или сульфидирования раствора.

Такая постановка задачи представляет интерес применительно к разработке методики расчета разделения релаксационным методом.

Поэтому, были разработаны алгоритм и программа расчета равновесного состояния в следующей постановке: методом последовательных приближений искать значения степеней сульфидирования и карбонизации, отвечающие заданным значениям параметров состояния (давления и температура процесса, концентрации извлекаемых компонентов в газовой или паровой фазах).

При построении итерационной процедуры могут быть использованы приемы, апробированные при отладке методики расчета разделения методом релаксации.

Алгоритм расчета построен следующим образом. По заданным значениям параметров состояния рассчитывается суммарное парциальное давление кислых компонентов в газовой фазе:

$$P_{H_2S} + P_{CO_2} = P_{\text{общ}} - P_{H_2O} - P_{\text{общ}} \sum Y_i \quad (9)$$

$P_{\text{общ}}$ - общее давление системы;

Y_i - концентрация инертных компонентов – углеводородов (мольн.доли)

P_{H_2O} - парциальное давление паров воды;

$$\lg P_{H_2O} = 8.65 + \lg \frac{\rho - N \cdot M_{ЭА}}{N \cdot M_{H_2O} \cdot M_{ЭА} + \rho} \quad (10)$$

Где $M(ЭА)$ и $M(H_2O)$ – молекулярные массы этаноламина и воды.

Затем задается соотношение β и при его фиксированном значении по известным значениям P_{H_2S} и P_{CO_2} находятся равновесные степени сульфидирования и карбонизации. После этого находится улучшенное значение β в двух последовательных итерациях с заданной точностью.

Необходимо отметить, что наибольшее влияние на результаты расчетов оказывает степень сульфи-

дирования (карбонизации) растворов, поэтому, была определена зависимость:

$$A = f(x) \quad (11)$$

Представленная в виде полинома:

$$A = a_0 + a_1x + a_2x^2 \quad (12)$$

При совместном присутствии сероводорода и диоксида углерода в очищаемом газе корректирующий множитель представлен зависимостью:

$$A' = f(\sum X_i) \quad (13)$$

$$\sum X_i = X_{H_2S} + X_{CO_2} \quad (14)$$

Результаты обработки данных методом наименьших квадратов позволили определить коэффициенты полиномов.

Корректирующие множители A_i , позволяют скорректировать полученные расчетные результаты и составляют основу разрабатываемой методики расчета упругости насыщенных паров над растворами аминов.

Заключение

В результате исследования равновесных свойств системы H_2S - CO_2 -МЭА была предпринята попытка определить и составить методику расчета термодинамического равновесия. Одним из преимуществ разрабатываемой программы вычисления парциальных давлений, является расширенный диапазон изменения степеней карбонизации, сульфидирования растворов этаноламинов и концентраций H_2S и CO_2 .

В целом можно говорить, что методика расчета равновесия удовлетворительно согласуется с экспериментальным материалом.

Решение поставленной задачи, необходимо для дальнейшей оптимизации и моделирования по заданному алгоритму процесса очистки углеводородов растворами этаноламинов, как локальных стадий процесса, так и глобальной оптимизации всего технологического процесса в целом.

Все это играет важную роль для развития математического моделирования круговых процессов очистки углеводородного сырья. В свою очередь, полученные результаты могут быть использованы для дальнейшего, более углубленного изучения процессов очистки углеводородов.

Список литературы

1. Мазгаров А.М. Технология очистки попутного нефтяного газа от сероводорода Учебное пособие / О.М. Корнетова. – Казань: Казан. Ун-т., 2015. – 70 с.
2. Soave G. Equilibrium constants from a modified Redlich–Kwong equation of state / G. Soave // Chemical engineering science. – 1972. – № 27. – P. 1197–1203.
3. Peng D.Y. A new two-constant equation of state /

D.Y. Peng, D.B. Robinson // Industrial and engineering chemistry: fundamentals. – 1976. – № 15. – P. 59–64.

4. Kent R.L. Better data for amine treating / R.L. Kent, B. Eisenberg // Hydrocarbon process. – 1976. – № 55 (2). – P. 87.

5. Li Y.G. Correlation and prediction of the solubility of carbondioxide in a mixed alkanolamine solution / Y.G. Li, A.E. Mather // Industrial and engineering chemistry research. – 1994. – № 33. – P. 2006–2015.

6. Анучин К.М., Мирошеченко Д.А., Возможность применения метилдиэтанолamina, активированного пиперазином в качестве абсорбента для глубокой очистки газа от CO₂, на основании расчетов в программе Petro-SIM.

ИССЛЕДОВАНИЕ МЕТОДОВ ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ УРАВНЕНИЙ КРУГОВОЙ ОГРАНИЧЕННОЙ ЗАДАЧИ ТРЁХ ТЕЛ

Гафурова Д.Р.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе изучаются методы интегрирования уравнений круговой ограниченной задачи трёх тел. Эта задача имеет один первый интеграл. Его значение должно быть постоянным при расчете решений во времени, однако без дополнительного контроля ошибок численного интегрирования оно отклоняется от исходного значения. В работе предлагается метод интегрирования, позволяющий оставить константу Якоби неизменной и повысить точность вычисления решений. Эффективность этого метода демонстрируется на примере периодической гало орбиты.

Введение

Круговая ограниченная задача трех тел [1] описывает движение малого тела в гравитационном поле, создаваемыми двумя массивными телами, движущимися по круговым орбитам вокруг общего центра масс. В данной задаче нет общего аналитического решения, кроме пяти стационарных точек, называемых точками либрации. В связи с этим для расчета орбит применяются численные методы. Особое внимание уделяется исследованию орбит около точек либрации, так как они используются на начальных этапах планирования космических миссий в качестве приближения движения космического аппарата.

С помощью метода дифференциальных коррекций в работах [2], [3] были рассчитаны семейства периодических гало орбит и n -периодических орбит, бифурцирующих из них. Этот метод был применен для построения основных семейств периодических орбит около точек либрации в [4], [5]. В работах [6]–[8] представлены численные методы, позволяющие

рассчитывать квазипериодические решения около точек либрации.

Круговая ограниченная задача трёх тел обладает одним первым интегралом, который называется константой Якоби или интегралом относительной энергии. Существование в данной задаче первого интеграла может быть использовано для уменьшения размерности системы. Эта идея реализуется в [9] для поиска начальных условий периодических орбит и в [10] для визуализации и анализа траекторий.

Из определения первого интеграла следует, что этот параметр не должен меняться при расчете решений во времени. Тем не менее, при вычислении орбит с помощью численных методов значение константы Якоби может отклоняться от исходного. Для решения проблемы сохранения значения первого интеграла в работах [7], [8] предлагается периодически прибавлять к вектору состояния коррекционный импульс, минимизирующий изменение этого параметра. При таком подходе отклонение константы Якоби от начального значения заметно уменьшается, но всё равно присутствует.

В данной работе предлагается метод интегрирования уравнений круговой ограниченной задачи трёх тел, учитывающий особенности этой задачи и сохраняющий константу Якоби. В нём производится уменьшение размерности системы за счёт первого интеграла. Использование нового метода интегрирования позволит повысить точность расчета орбит и их матриц перехода, собственные значения которых играют большую роль в исследовании решений.

Круговая ограниченная задача трёх тел

Для описания движения материальной точки в круговой ограниченной задаче трёх тел вводится вращающаяся система координат, начало которой совпадает с барицентром массивных тел P_1, P_2 . Ось Ox направлена в сторону меньшего из двух тел, направление оси Oz совпадает с вектором угловой скорости P_1, P_2 , а ось Oy дополняет Ox, Oz до правой тройки. Расстояние между массивными телами принимается равным единице. В этой системе тела P_1, P_2 покоятся, а динамика малого тела описывается следующей системой дифференциальных уравнений

$$\begin{cases} \dot{x} = v_x \\ \dot{y} = v_y \\ \dot{z} = v_z \\ \dot{v}_x = 2\dot{y} + x - \frac{\mu_1(x+\mu_2)}{r_1^3} - \frac{\mu_2(x-\mu_1)}{r_2^3} \\ \dot{v}_y = -2\dot{x} + y - \frac{\mu_1 y}{r_1^3} - \frac{\mu_2 y}{r_2^3} \\ \dot{v}_z = -\frac{\mu_1 z}{r_1^3} - \frac{\mu_2 z}{r_2^3}, \end{cases} \quad (1)$$

где $r_1 = \sqrt{(x+\mu_2)^2 + y^2 + z^2}$, $r_2 = \sqrt{(x-\mu_1)^2 + y^2 + z^2}$, $\mu_1 = \frac{m_1}{m_1+m_2}$, $\mu_2 = \frac{m_2}{m_1+m_2}$.

Соответственно, механическое состояние материальной точки задаётся шестимерным вектором $(x, y, z, v_x, v_y, v_z)^T$.

Описание метода интегрирования уравнений

Известно, что круговая ограниченная задача трёх тел обладает первым интегралом:

$$C_J = x^2 + y^2 + 2 \left(\frac{\mu_1}{r_1} + \frac{\mu_2}{r_2} \right) - v^2. \quad (2)$$

Вводится замена переменных

$$\begin{cases} v_x = \rho \cdot \cos \varphi \\ v_y = \rho \cdot \sin \varphi, \end{cases} \quad (3)$$

в которой константа Якоби принимает следующий вид:

$$C_J = x^2 + y^2 + 2 \left(\frac{\mu_1}{r_1} + \frac{\mu_2}{r_2} \right) - \rho^2 - v_z^2. \quad (4)$$

При фиксированном значении константы Якоби введенная переменная ρ может быть представлена функцией:

$$\rho(x, y, z, v_z) = \sqrt{x^2 + y^2 + 2 \left(\frac{\mu_1}{r_1} + \frac{\mu_2}{r_2} \right) - v_z^2 - C_J}. \quad (5)$$

Таким образом, наличие первого интеграла допускает уменьшение размерности системы, описывающей движение малого тела

$$\begin{cases} \dot{x} = \rho(x, y, z, v_z) \cdot \cos \varphi \\ \dot{y} = \rho(x, y, z, v_z) \cdot \sin \varphi \\ \dot{z} = v_z \\ \dot{\varphi} = \frac{Ay \cos \varphi - B \sin \varphi}{\rho(x, y, z, v_z)} - 2 \\ \dot{v}_z = -\frac{\mu_1 y}{r_1^3} - \frac{\mu_2 y}{r_2^3}, \end{cases} \quad (6)$$

где $A = \left(1 - \frac{\mu_1}{r_1^3} - \frac{\mu_2}{r_2^3} \right)$,

$$B = \left(x - \frac{\mu_1(x+\mu_2)}{r_1^3} - \frac{\mu_2(x-\mu_1)}{r_2^3} \right).$$

Система дифференциальных уравнений (6) интегрируется явным методом Рунге-Кутты восьмого порядка. По начальным значениям $(x_0, y_0, z_0, \rho_0, \varphi_0, v_{z0})^T$ рассчитывается константа Якоби, а затем на каждом шаге интегрирования ρ пересчитывается с учетом текущих значений x, y, z, v_z и исходного значения C_J . За счет этого константа Якоби одинакова во всех точках получаемой траектории.

Представленный метод интегрирования позволяет рассчитывать орбиты, которые ни в одной точке не ортогональны плоскости вращения массивных тел. Для большинства известных периодических семейств это условие выполняется. При расчете новых семейств, появление таких орбит определяется по стремлению минимального значения ρ к нулю. В этом случае необходимо использовать другие методы интегрирования. Метод реализован на языке программирования Python с использованием программного модуля Orbipy [11].

Описание алгоритма коррекционных импульсов

Интегрирование уравнений круговой ограниченной задачи трёх тел сопровождается накоплением численных ошибок, из-за которых при расчете периодической орбиты на долгий промежуток времени её траектория может отдаляться от истинного решения. Один из способов, позволяющих решить эту проблему и рассчитывать орбиты до длительного момента времени, заключается в добавлении к вектору состояния коррекционных импульсов $\Delta \vec{v} = (\Delta v_x, \Delta v_y, \Delta v_z)^T$. Эти коррекционные импульсы являются виртуальными и компенсируют ошибки численного интегрирования.

В данной работе коррекционные импульсы прибавлялись к вектору состояния материальной точки в моменты времени кратные $\frac{T}{2}$, где T - период интегрируемого решения. Направление \vec{u} вектора $\Delta \vec{v} = \Delta v \vec{u}$ совпадало с вектором неустойчивого направления. Это такое направление, что после сообщения малому телу импульса \vec{u} возмущенная траектория максимально отдаляется от исходной орбиты через промежуток времени T .

Поиск величины коррекционного импульса Δv производился с использованием сечения Пуанкаре интегрируемой периодической орбиты. Под сечением Пуанкаре орбиты некоторой плоскостью имеется в виду множество её точек пересечения с этой плоскостью. Сечение Пуанкаре периодической орбиты является конечным набором точек. Величина Δv выбиралась такой, что после прибавления к вектору состояния коррекционного импульса $\Delta \vec{v} = \Delta v \vec{u}$ и расчета этого состояния на период, расстояние между точками в сечении Пуанкаре новой траектории и точками в заданном сечении Пуанкаре принимало наименьшее значение.

Исследование точности предложенного метода

Рисунок 1 представляет график изменения константы Якоби во времени для гало орбиты около точки либрации L1 в системе Земля-Луна. Начальное состояние орбиты этой орбиты

$$\vec{X}_0 = \left(378847.42 \text{ км}, 0, 21589.21 \text{ км}, 0, -0.63341 \frac{\text{км}}{\text{с}}, 0 \right)^T.$$

Орбита была проинтегрирована двумя способами. В первом случае численное решение было найдено путём интегрирования системы уравнений (1) методом Рунге-Кутты восьмого порядка. Значение константы Якоби орбиты, построенной этим способом, отображено на графике зеленым цветом. Во втором случае применялся предложенный в работе метод; интеграл Якоби, относящийся к этому примеру, показан красным цветом на графике. При расчете решения с нулевого момента времени до момента равного 131 дню первым интегратором значение интеграла Якоби варьировалось в порядке 10^{-12} , в то время как представленный метод интегрирования позволил сохранить этот параметр неизменным.

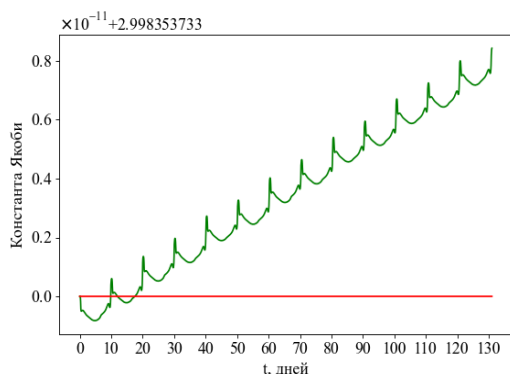


Рис. 1. Изменение константы Якоби при расчете гало орбиты методом, не учитывающим интеграл Якоби (зеленая линия), и предложенным в работе способом (красная линия)

Для сравнения точности работы этих интеграторов периодическая гало орбита с тем же начальным условием была рассчитана до момента времени равного 75 периодам (примерно 755 дней). При вычислении орбиты также применялся алгоритм коррекционных импульсов. На рисунке 2 показаны полученные сравниваемыми интеграторами траектории движения малого тела и их сечения Пуанкаре плоскостью $y = 0$.

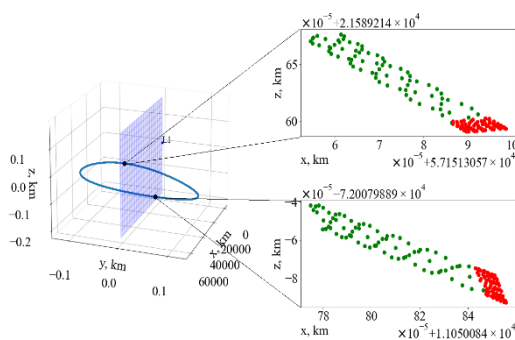


Рис. 2. Гало орбита около точки либрации L1 (слева) и сечение Пуанкаре при интегрировании двумя разными способами (справа)

Зеленые точки соответствуют сечению Пуанкаре орбиты, рассчитанной с помощью первого способа интегрирования, а красные – сечению Пуанкаре орбиты, найденной предложенным методом. Известно, что сечение Пуанкаре гало орбиты плоскостью $y = 0$ состоит из двух точек, то есть рассчитанные траектории пересекают плоскость $y = 0$ по два раза за один период. Из-за ошибок численного интегрирования на каждом новом периоде пара точек пересечения может отличаться от предыдущих точек пересечения плоскости. Тем не менее, при расчете орбиты представленным методом разброс точек в сечении Пуанкаре примерно в 9 раз меньше, чем при интегрировании альтернативным способом. Таким образом, предложенный в работе метод дал возможность уменьшить ошибку при расчете гало орбиты.

Заключение

В работе представлен метод интегрирования уравнений круговой ограниченной задачи трёх тел, сохраняющий значение константы Якоби. Использование этого метода повышает точность расчета орбит, что показано на примере построения гало орбиты. Дальнейшая работа будет направлена на расчет новых семейств периодических орбит вокруг коллинеарных точек либрации круговой ограниченной задачи трех тел.

Список литературы

1. Мюррей К., Дермотт С. Динамика Солнечной системы / Пер. с англ. под ред. И. И. Шевченко. - М.: ФИЗМАТЛИТ, 2010. - 588 с. - ISBN 978-5-9221-1121-8.
2. K. Connor Howell, "Three-dimensional, periodic, halo orbits," *Celest. Mech.*, vol. 32, no. 1, pp. 53–71, Jan. 1984, doi: 10.1007/BF01358403.
3. E. T. Campbell, "Bifurcations from families of periodic solutions in the circular restricted problem with application to trajectory design," Purdue University ProQuest Dissertations Publishing, 1999.
4. D. C. Folta, N. Bosanac, D. Guzzetti, and K. C. Howell, "An earth-moon system trajectory design reference catalog," *Adv. Astronaut. Sci.*, vol. 153, pp. 191–210, 2015, doi: 10.1016/j.actaastro.2014.07.037.
5. L. Bury, J. McMahon, and M. Lo, "A study of periodic orbits near Europa," *Celest. Mech. Dyn. Astron.*, vol. 134, no. 3, Jun. 2022, doi: 10.1007/s10569-022-10076-6.
6. K. C. Howell and H. J. Pernicka, "Numerical determination of Lissajous trajectories in the restricted three-body problem," *Celest. Mech.*, vol. 41, no. 1–4, pp. 107–124, Mar. 1987, doi: 10.1007/BF01238756.
7. Ren Y., Shan J. A novel algorithm for generating libration point orbits about the collinear points // *Celestial Mechanics and Dynamical Astronomy*. – 2014. – T. 120. – С. 57-75.
8. Aksenov S., Bober S., Guskova M. Mapping of initial conditions for libration point orbits // *Advances in Space Research*. – 2021. – T. 68. – №.6. – С. 2501-2514.
9. Kumar V., Kushvah B. S. Computation of periodic orbits around L 1 and L 2 using pso technique // *Astronomy Reports*. – 2020. – T. 64. – №. 1. – С. 82-93.
10. A. F. Haapala and K. C. Howell, "Representations of higher-dimensional Poincaré maps with applications to spacecraft trajectory design," *Acta Astronaut.*, vol. 96, pp. 23–41, Mar. 2014, doi: 10.1016/j.actaastro.2013.11.019.
11. Bober Stanislav, Aksenov Sergey, & Guskova Mariia. (2023, January 20). OrbiPy (Version 0.2.4). Zenodo. <http://doi.org/10.5281/zenodo.4301584>

ПОСТРОЕНИЕ МОДЕЛИ ДЕФОРМАЦИОННОГО ПОВЕДЕНИЯ СВЕРХПЛАСТИЧЕСКИХ МАТЕРИАЛОВ

Миколаенко В.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ*

Аннотация

Листовые заготовки из сверхпластичных материалов используются в технологиях сверхпластической газовой формовки для производства тонкостенных изделий сложной геометрической формы. При реализации такой технологии необходимо корректно задать режим давления и температуры, чтобы заготовка вошла в состояние сверхпластичности, при котором возможно её бездефектное формоизменение при больших деформациях. Режим давления и температуры устанавливаются исходя из модели деформационного поведения материала, которая связывает степень и скорость деформации с напряжением. Такую модель строят так, чтобы она аппроксимировала серию кривых «напряжение-деформация», полученных экспериментально. Как правило, такая модель содержит множество параметров. В данной работе предложен метод построения такой модели, позволяющий минимизировать количество параметров при заданной точности модели. Разработанный метод был опробован на примере построения свойств сплава системы Al-Mg-Fe-Ni.

Введение

Сверхпластичность – это состояние материала, которое при бездефектном формоизменении допускает деформации, значительно превышающие максимально возможные для этого материала в обычном состоянии [1]. При сверхпластической газовой формовке листовых заготовок используются режимы давления и температуры, учитывающие особенности формоизменения деформируемого сплава [2,3]. Данные особенности заложены в модели деформационного поведения материала, связывающей напряжения с деформациями и скоростями деформации [4]. Существует множество экспериментальных методик, позволяющих получить такую модель [5–7]. Основной экспериментальной методикой является испытание плоских образцов на растяжение при постоянной скорости деформации [8]. Процедура проведения такого испытания, геометрические особенности установки и образца, а также интерпретации результатов испытаний описаны в стандарте ASTM-E2448 [9]. Интерпретация результатов испытания по экспериментальным данным заключается в построении кривой «напряжение-деформация» по измеренным за время испытания силам. По серии таких кривых, полученных из испытаний с разными постоянными скоростями деформации, проводят построение модели деформационного поведения материала, аппроксимирующей полученные зависимости.

При интерпретации экспериментальных данных существует проблема неоднородности распределения скоростей деформации в образце, что может приводить к ошибкам в оценке напряжений порядка 20%. Данная проблема обсуждалась во многих работах [10–13]. Для решения данной проблемы в работе [14] была предложена итерационная процедура обратного анализа, на каждом шаге которой проводится конечно-элементное моделирование, позволяющее учесть особенности течения материала в образце. Сначала конечно-элементное моделирование проводится при использовании свойств материала, построенных на экспериментальных данных, но после на каждой итерации модель деформационного поведения уточняется, при этом используются результаты конечно-элементного моделирования. На каждом шаге итерационной процедуры после проведения моделирования проводится построение уточнённой модели деформационного поведения материала.

В данной работе решается задача построения модели деформационного поведения материала по серии кривых «напряжение-деформация», полученных при разных постоянных скоростях деформации. Эта задача заключается в построении начального приближения и его дальнейшем уточнении.

Процедура построения модели деформационного поведения материала

Модель деформационного поведения материала представляет собой зависимость, связывающую напряжения σ с деформациями ε и скоростями деформации $\dot{\varepsilon}$:

$$\sigma = \sigma(\varepsilon, \dot{\varepsilon}). \quad (1)$$

Процедура построения начального приближения совпадает с процедурой, опубликованной в работах [15,16]. При этом используются уравнения состояния деформируемой среды:

$$\dot{\varepsilon} = A_1 \sigma^n, \quad (2)$$

$$\dot{\varepsilon} = A_2 \exp(b\sigma), \quad (3)$$

$$\dot{\varepsilon} = A_3 (\sinh(a\sigma))^\gamma, \quad (4)$$

где $A_1, A_2, A_3, b, n, \gamma, a$ – коэффициенты, описывающие модель материала.

После логарифмирования уравнений (2) и (3) с помощью линейной регрессии вычисляют коэффициенты материала n и b соответственно:

$$\ln \dot{\varepsilon} = \ln A_1 + n \ln \sigma \Rightarrow n, \quad (5)$$

$$\ln \dot{\varepsilon} = \ln A_2 + b\sigma \Rightarrow b. \quad (6)$$

На основе вычисленных коэффициентов из уравнений (5) и (6) определяют коэффициент a из уравнения (4):

$$a = b/n. \quad (7)$$

После логарифмирования уравнения (4) и подстановки в него коэффициента a из уравнения (7) с помощью линейной регрессии вычисляют коэффици-

енты материала A_3 и γ , при этом удобно ввести обозначение $\beta = \ln A_3$:

$$\ln \varepsilon = \ln A_3 + \gamma \ln \sinh\left(\frac{\beta}{n} \sigma\right) \Rightarrow \beta, \gamma. \quad (8)$$

Для построения непрерывной зависимости проводят аппроксимацию полученного набора коэффициентов материала α_i , β_i и γ_i полиномами пятой степени:

$$\alpha(\varepsilon) = K_1 + K_2\varepsilon + K_3\varepsilon^2 + K_4\varepsilon^3 + K_5\varepsilon^4 + K_6\varepsilon^5, \quad (9)$$

$$\beta(\varepsilon) = K_7 + K_8\varepsilon + K_9\varepsilon^2 + K_{10}\varepsilon^3 + K_{11}\varepsilon^4 + K_{12}\varepsilon^5, \quad (10)$$

$$\gamma(\varepsilon) = K_{13} + K_{14}\varepsilon + K_{15}\varepsilon^2 + K_{16}\varepsilon^3 + K_{17}\varepsilon^4 + K_{18}\varepsilon^5. \quad (11)$$

Модель деформационного поведения материала из уравнения (1) получают, выражая напряжения σ из уравнения (4) и используя в качестве коэффициентов материала непрерывные зависимости из уравнений (9), (10) и (11):

$$\sigma(\varepsilon, \dot{\varepsilon}) = \sigma(\alpha(\varepsilon), \beta(\varepsilon), \gamma(\varepsilon), \dot{\varepsilon}) = \frac{1}{\alpha} \ln \left((\dot{\varepsilon}/e^\beta)^{1/\gamma} + \sqrt{1 + (\dot{\varepsilon}/e^\beta)^{2/\gamma}} \right). \quad (12)$$

Полученное начальное приближение обладает следующими недостатками: оно не обеспечивает минимума среднеквадратичного отклонения модели от экспериментальных данных, а также включает в себя 18 параметров, что не позволяет поставить задачу минимизации в 18-мерном пространстве.

При уточнении модели деформационного поведения материала предлагается для каждого значения степеней деформации ε_i скорректировать соответствующий ему набор коэффициентов материала $\alpha_i, \beta_i, \gamma_i$ с помощью метода Нелдера-Мида, минимизируя среднеквадратичное отклонение рассчитываемых напряжений от экспериментальных:

$$d_i = \sum_j^M \left(\sigma_{i,j}^{\text{эксп}} - \sigma(\alpha_i, \beta_i, \gamma_i, \dot{\varepsilon}_j) \right)^2, \quad (13)$$

где $\sigma_{i,j}^{\text{эксп}}$ – значение напряжений, соответствующих деформации ε_i у кривой, полученной из испытания со скоростью деформации $\dot{\varepsilon}_j$, M – количество испытаний с разными скоростями деформации. Используя в качестве начального приближения рассчитанные ранее коэффициенты материала α_i , β_i и γ_i , получим набор скорректированных коэффициентов $\tilde{\alpha}_i$, $\tilde{\beta}_i$ и $\tilde{\gamma}_i$. Каждый из наборов коэффициентов материала, соответствующий ε_i деформации, скорректирован по отдельности. Из-за отсутствия явного локального минимума результаты получаются зашумленными, что приводит к большим ошибкам при аппроксимации.

С целью уменьшения разброса коэффициентов предлагается последовательно строить аппроксимации $\tilde{\alpha}(\varepsilon)$, $\tilde{\beta}(\varepsilon)$ и $\tilde{\gamma}(\varepsilon)$, проводя последовательный

запуск методов оптимизации на оставшихся неаппроксимированных коэффициентах. Причем аппроксимации могут проводиться в любом порядке, используя полиномы любой степени. Полученные в результате аппроксимации будут включать в себя параметры $K_{\alpha 0}, \dots, K_{\alpha n}, K_{\beta 0}, \dots, K_{\beta m}, K_{\gamma 0}, \dots, K_{\gamma k}$, описывающие скорректированную модель деформационного поведения материала $\tilde{\sigma}(\varepsilon, \dot{\varepsilon})$:

$$\tilde{\alpha}(\varepsilon) = K_{\alpha 0} + K_{\alpha 1}\varepsilon + K_{\alpha 2}\varepsilon^2 + \dots + K_{\alpha n}\varepsilon^n, \quad (14)$$

$$\tilde{\beta}(\varepsilon) = K_{\beta 0} + K_{\beta 1}\varepsilon + K_{\beta 2}\varepsilon^2 + \dots + K_{\beta m}\varepsilon^m, \quad (15)$$

$$\tilde{\gamma}(\varepsilon) = K_{\gamma 0} + K_{\gamma 1}\varepsilon + K_{\gamma 2}\varepsilon^2 + \dots + K_{\gamma k}\varepsilon^k, \quad (16)$$

где n , m и k , – степени соответствующих полиномов.

Множество построенных моделей получено из-за разной последовательности аппроксимации коэффициентов материала и из-за разного количества параметров, начиная от трёх параметров, что соответствует аппроксимации коэффициентов тремя полиномами нулевой степени. Среди данного множества отбирается модель, удовлетворяющая требуемому количеству параметров модели или требуемой точности.

Обработка экспериментальных данных

Предложенный метод построения модели деформационного поведения материала был автоматизирован с помощью Python скриптов и опробован на экспериментальных данных, полученных из испытаний на растяжение плоских образцов из сплава системы Al-Mg-Fe-Ni [17]. Серия испытаний проводилась при температуре 460°C и при следующих постоянных скоростях деформации: 0.002c^{-1} , 0.005c^{-1} , 0.01c^{-1} , 0.03c^{-1} и 0.06c^{-1} . Геометрические параметры образца и установки, проводящей испытания, демонстрируются на рисунке 1.

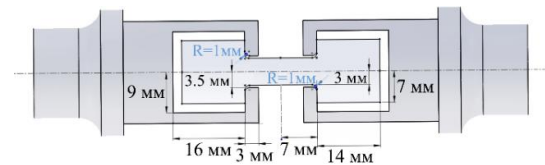


Рис. 1. Геометрические параметры образца и установки

Начальное приближение модели деформационного поведения материала $\sigma(\varepsilon_i, \dot{\varepsilon}_j)$ изображено на рисунке 2 (непрерывными линиями) в сравнении с экспериментальными данными $\sigma_{i,j}^{\text{эксп}}$ (пунктирными линиями с маркерами).

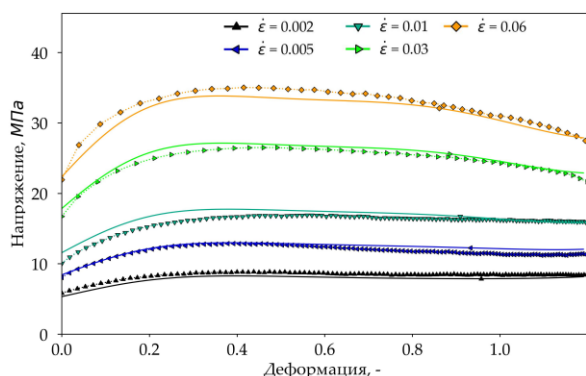


Рис. 2. Сравнение кривых напряжения-деформации, полученных из экспериментальных данных, с начальным приближением свойств материала

Затем было получено множество уточнённых моделей с разным количеством параметров при разной последовательности аппроксимаций коэффициентов материала. Далее для каждой из последовательностей и для каждого из параметров среди имеющихся моделей были рассчитаны среднеквадратичные отклонения от экспериментальных данных и выбраны лучшие. Результат такого отбора изображён на рисунке 3, где в легенде приведены последовательности аппроксимаций, по оси абсцисс количество параметров модели, а по оси ординат среднеквадратичные отклонения напряжений. Также на рисунке 3 пунктирной линией отображена точность, получаемая с помощью начального приближения.

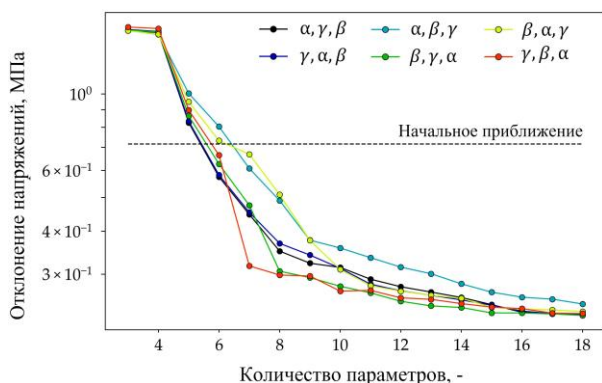


Рис. 3. Отклонения скорректированных моделей от кривых напряжения-деформации, полученных из экспериментальных данных

Исходя из полученных результатов видно, что у скорректированной модели с 7 параметрами удаётся повысить точность более, чем в два раза, в сравнении с начальным приближением, при построении которого использовалось 18 параметров. Полученная скорректированная модель построена при последовательном запуске методов оптимизации, при котором аппроксимации коэффициентов материала проводились в следующем порядке: γ , β , α . В таблице 1 приведены параметры такой модели.

Таблица 1. Параметры модели

$\alpha(\epsilon)$		$\beta(\epsilon)$		$\gamma(\epsilon)$	
$K_{\alpha 0}$	0,00426	$K_{\beta 0}$	2,3627	$K_{\gamma 0}$	2,57385
$K_{\alpha 1}$	0,00153			$K_{\gamma 1}$	-7,12
$K_{\alpha 2}$	0,0232			$K_{\gamma 2}$	1,7672

В таблице 2 приведено сравнение среднеквадратичных отклонений экспериментальных данных $\sigma_{i,j}^{\text{эксп}}$ от начального приближения $\sigma(\epsilon, \dot{\epsilon})$ и от скорректированной модели $\tilde{\sigma}(\epsilon, \dot{\epsilon})$, параметры которой описаны в таблице 1.

Таблица 2. Отклонение кривых напряжения деформации из экспериментальных данных от полученных моделей

Скорости деформации, с^{-1}	Отклонение $\sigma_{i,j}^{\text{эксп}}$ от $\sigma(\epsilon, \dot{\epsilon})$, МПа	Отклонение $\sigma_{i,j}^{\text{эксп}}$ от $\tilde{\sigma}(\epsilon, \dot{\epsilon})$, МПа
$\dot{\epsilon}_1 = 0.002 \text{с}^{-1}$	0,556	0,277
$\dot{\epsilon}_2 = 0.005 \text{с}^{-1}$	0,478	0,43
$\dot{\epsilon}_3 = 0.01 \text{с}^{-1}$	0,878	0,114
$\dot{\epsilon}_4 = 0.03 \text{с}^{-1}$	0,617	0,338
$\dot{\epsilon}_5 = 0.06 \text{с}^{-1}$	1,047	0,355
Среднее:	0,715	0,317

Из таблицы 2 видно, что точность удалось повысить более, чем в два раза, при меньшем количестве параметров. На рисунке 4 представлено сравнение экспериментальных данных $\sigma_{i,j}^{\text{эксп}}$ (пунктирные линии с маркерами) со скорректированной моделью $\tilde{\sigma}(\epsilon, \dot{\epsilon})$ (непрерывные линии), параметры которой описаны в таблице 1.

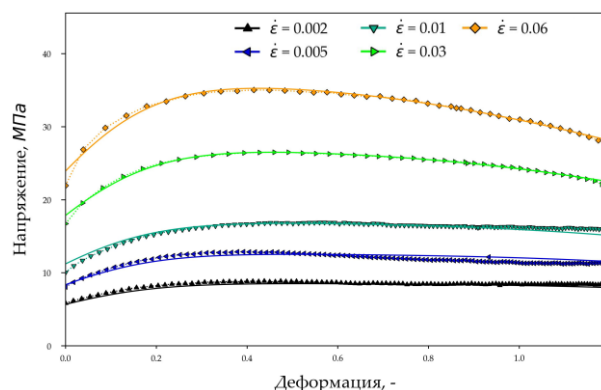


Рис. 4. Сравнение кривых напряжения-деформации, полученных из экспериментальных данных, со скорректированной моделью

По рисунку 4 видно, что кривые, полученные с помощью скорректированной модели, согласуются с экспериментальными данными лучше, чем это было на рисунке 2, где использовалось начальное приближение.

Заключение

В данной работе предложен метод построения модели деформационного поведения материала по серии кривых «напряжение-деформация», позволяющий минимизировать количество параметров, требуемых для описания свойств материала при заданных требованиях на точность модели. Разработанная процедура была автоматизирована с помощью Python скриптов и опробована на экспериментальных данных, полученных из серии испытаний на одноосное растяжение плоских образцов сплава системы Al-Mg-Fe-Ni. В результате была построена модель, которая в сравнении с начальным приближением обеспечивает более чем двукратное повышение точности при почти в три раза меньшем количестве параметров. Результаты исследования будут использоваться в дальнейшем при доработке методики коррекции экспериментальных данных из работы [14].

Благодарность

Автор выражает благодарность лаборатории «Ультрамелкозернистых металлических материалов» НИТУ «МИСиС» за предоставление экспериментальных данных.

Список литературы

1. Sherby O. D., Wadsworth J. Superplasticity-Recent advances and future directions. - 1989. - *Prog. Mater. Sci.*
2. Jarrar F., Sorgente D., Aksenov S., Enikeev F. On the challenges and prospects of the superplastic forming process. - 2018. *Materials Science Forum* vol 941 MSF (Trans Tech Publications Ltd) pp 2343–8
3. Aksenov S. A. S. A. S. A., Kolesnikov A. V. A. V., Mikhaylovskaya A. V. A. V. Design of a gas forming technology using the material constants obtained by tensile and free bulging testing. - 2016. - *J. Mater. Process. Technol.* - № 237. - с. 88–95.
4. Enikeev F. U., Kruglov A. A. An analysis of the superplastic forming of a thin circular diaphragm. - 1995. - *Int. J. Mech. Sci.*
5. Giuliano G., Franchitti S. On the evaluation of superplastic characteristics using the finite element method. - 2007. - *Int. J. Mach. Tools Manuf.* - № 47. - с. 471–6.
6. El-Morsy A., Akkus N., Manabe K., Nishimura H. Evaluation of superplastic characteristics of tubular materials by multi-tube bulge test. - 2006. - *Mater. Lett.*
7. Vulcan M., Siegert K., Banabic D. The influence of pulsating strain rates on the superplastic deformation behaviour of Al-alloy AA5083 investigated by means of cone test. - 2004. *Materials Science Forum*
8. Yasmeen T., Shao Z., Zhao L., Gao P., Lin J., Jiang J. Constitutive modeling for the simulation of the superplastic forming of TA15 titanium alloy. - 2019. - *Int. J. Mech. Sci.*
9. E2448-11 Standard Test Method for Determining the Superplastic Properties of Metallic Sheet Materials. - 2011. - *ASTM B. Stand.*
10. Aksenov S., Mikolaenko V. The effect of material properties on the accuracy of superplastic

tensile test. - 2020. - *Metals (Basel).*

11. Abu-Farha F., Nazzal M., Curtis R. Optimum Specimen Geometry for Accurate Tensile Testing of Superplastic Metallic Materials. - 2011. - *Exp. Mech.*

12. Bate P. S., Ridley N., Sotoudeh K. Effect of gauge length in superplastic tensile tests. - 2008. - *Mater. Sci. Technol.*

13. Миколаенко В. Анализ деформационного поведения экспериментального сплава системы AL-MG-FE-NI в условиях сверхпластичности. - 2021. *Материалы межвузовской научно-технической конференции студентов, аспирантов и молодых специалистов имени Е.В. Арменского* pp 39–42

14. Миколаенко В. Коррекция модели деформационного поведения сверхпластических материалов. - 2022. *Материалы межвузовской научно-технической конференции студентов, аспирантов и молодых специалистов имени Е.В. Арменского* pp 34–7

15. Mosleh A. O. A. O., Mikhaylovskaya A. V. A. V., Kotov A. D. A. D., Kwame J. S. J. S., Aksenov S. A. S. A. Superplasticity of Ti-6Al-4V titanium alloy: Microstructure evolution and constitutive modelling. - 2019. - *Materials (Basel).* - № 12.

16. Mosleh A. O., Mikhaylovskaya A. V., Kotov A. D., Portnoy V. K. Arrhenius-type constitutive equation model of superplastic deformation behaviour of titanium-based alloys. - 2018. *Defect and Diffusion Forum*

17. Kishchik A. A., Mikhaylovskaya A. V., Kotov A. D., Rofman O. V., Portnoy V. K. Al-Mg-Fe-Ni based alloy for high strain rate superplastic forming. - 2018. - *Mater. Sci. Eng. A.* - № 718. - с. 190–7.

ОБ АВТОМАТИЧЕСКОЙ ГЕНЕРАЦИИ СИ-КОДА НА ОСНОВЕ EVENT-В СПЕЦИФИКАЦИЙ

Зеленецкая Д.Н.¹, Пучков Д.В.², Десянин П.Н.³

¹*Национальный исследовательский университет «Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ,*

²*Московский физико-технический институт (национально исследовательский университет),
физтех-школа радиотехники и*

*компьютерных технологий,
³ООО «РусБИТех-Астра»*

Аннотация

Исследование посвящено разработке способа генерации кода на языке программирования Си (Си-кода) и соответствующих ему спецификаций на языке ACSL (ACSL-спецификаций), исходя из описания математической модели безопасности управления доступом и информационными потоками в операционной системе специального назначения (ОСЧН) Astra Linux (МРОСЛ ДП-модели), представленной на языке формального метода Event-

В. При этом основное внимание акцентировано на генерации Си-кода и ACSL-спецификаций механизмов и системных вызовов, выполняющих управление доступом в ядре ОСЧН, а также способе разработки программного решения для проверки корректности такой реализации МРОСЛ ДП-модели в Си-коде.

Введение

В результате решения задач обеспечения технологической независимости в области информационных технологий и информационной безопасности существенно возросли объемы разрабатываемого отечественного системного программного обеспечения (ПО), при этом становится все более актуальным достижение доверия к такому ПО, в том числе путем выполнения требований профильных нормативных документов регуляторов. В связи с этим в соответствии с формируемой с участием авторов методологией разработки безопасного системного ПО [1] одной из задач здесь является дальнейшее развитие и автоматическая верификация согласно ГОСТ Р 59453.1,2-2021 [2, 3] математической модели (МРОСЛ ДП-модели) [4], формально описывающей и задающей функциональность механизмов мандатного и дискреционного (ролевого) управления доступом и мандатного контроля целостности ОСЧН Astra Linux [5].

Постановка решаемой задачи

МРОСЛ ДП-модель кроме математического языка также была описана с помощью языка формального метода Event-B [6], что позволило автоматически верифицировать ее с помощью инструментального средства Rodin [7]. В результате было получено формализованное описание модели, которое было названо «верхнеуровневым». Следующим шагом достижения доверия стала проверка корректности реализации модели непосредственно в программном коде ОСЧН. На данный момент, для обнаружения уязвимостей, возникающих за счет нарушения логики управления доступом, применяются инструментальные средства для верификации кода [8], такое как Frama-C, поддерживающее формальный язык спецификаций ACSL [9]. Но в связи с большим объемом кода ОСЧН, частым выходом ее новых релизов, трудоемкостью верификации, авторами применяется иной способ решения задачи, который включает в себя поочередную автоматическую генерацию Си-кода на основе описания МРОСЛ ДП-модели на языке формального метода Event-B и автоматическую генерацию ACSL-спецификаций на основе сгенерированного кода. Этот способ также позволяет сравнивать ACSL-спецификации для сгенерированного и уже существующего кода.

Чтобы использовать рассматриваемый способ оказалось недостаточным абстрактного «верхнеуровневого» формализованного описания модели. Так как язык Event-B основан на теории

множеств и логике предикатов, а модель включает описание состояний, переходов между состояниями, называемых событиями (состоящими из переменных, условий и результатов выполнения), и свойствами (инвариантами), которые должны соблюдаться в каждом состоянии, то для достижения наибольшей точности перевода модели в Си-код целесообразно использовать ее описание, синтаксически более близкое к существующему Си-коду ОСЧН. Для этого потребовалось разработать еще одно формализованное описание модели (названное «низкоуровневым» или моделью функциональных спецификаций [8, 10]), которое соответствует системным вызовам ОСЧН, максимально приближенно к их реализации в программном коде.

Стоит также отметить, что такое «низкоуровневое» описание модели, оперируя терминами, присутствующими в таком же контексте в Си-коде соответствующих системных вызовов, увеличивает визуальное сходство сгенерированного кода с существующим и дает возможность сравнения сгенерированного кода с реализацией аналогичных функций в Си-коде ядра ОСЧН

Наглядным примером этого является определение флагов прав доступа на чтение (UREAD), запись (UWRITE) и выполнение (UEXECUTE) в «низкоуровневом» описании модели, которое визуально соответствует применению аналогичных флагов в существующем Си-коде механизма управления доступом ОСЧН (листинг 1).

```
@S_IRUSR_type: S_IRUSR = UREAD
@S_IWUSR_type: S_IWUSR = UWRITE
@S_IXUSR_type: S_IXUSR = UEXECUTE
```

Листинг 1. Пример определения флагов прав доступа в модели.

Реализация поставленной задачи

Для программной реализации рассматриваемого способа были проанализированы компоненты «низкоуровневого» описания модели, а именно: контекст (context), состоящий из множеств (sets), константных значений (constants) и аксиом (axioms), и машина (machine), содержащая в себе переменные машины (variables), инварианты событий (invariants) и сами события (events) [8]. Для того, чтобы генерируемый Си-код больше соответствовал коду механизма управления доступом в ядре ОСЧН, были реализованы следующие основные решения:

- Для множеств (sets) инициализация в Си-коде происходит при помощи типов данных, которые заранее задаются через зарезервированное ключевое слово определения типов в «низкоуровневом» формализованном описании модели;

- Для известных констант (constants) инициализация в Си-коде происходит при помощи директив препроцессора (аналогично осуществляется инициализация констант и в существующем коде ядра ОСЧН). Для остальных случаев (будем называть их составные переменные) инициализация происходит в разделе аксиомы

(axioms), с учетом конкретики каждой константной переменной;

- Для аксиоматических выражений (axioms) инициализирование составных переменных в Си-коде производится в зависимости от соответствующих условий: либо вводится перечисление, либо осуществляется инициализирование при помощи директив препроцессора и зарезервированного ключевого слова определения типов (использованного для инициализации множеств sets). Для остальных конструкций задается соответствующая структура данных в Си-коде (пользовательская структура).

Для перевода соответствующих системным вызовам событий (events), в которых осуществляются проверки принадлежности параметров события некоторым множествам, первоначально использовались n-мерные массивы и структуры (например, для проверки, что процесс обладает множеством открытых файловых дескрипторов: $FDFlags \in FDs \rightarrow P(OPEN_FLAGS)$, использовался 2-мерный массив вида $FDFlags[int][int]$). Однако, апробация этого подхода при переводе в Си-код события `create_hard_link`, соответствующего системному вызову `link`, создающему ссылку на файл или каталог (выбор именно этого события был сделан в связи с существенным сходством описания события и функционалом системного вызова, а также тем, что ACSL-спецификации для сгенерированного кода были близки к ранее разработанным ACSL-спецификациям существующего Си-кода системного вызова `link` и их было удобно сопоставлять), было выявлено, что данное решение не подходит для дальнейшей работы. Это объясняется тем, что специфика формального языка ACSL не позволяет верифицировать код, включающий сравнение элементов определенных ранее пользовательских структур, а также включающий, вложенные циклы, которые возникают при использовании n-мерных массивов. Поэтому для перевода соответствующих системным вызовам событий (events) было принято решение использовать хеш-таблицы (в приведенном примере это хеш-таблица `struct HT FDFlags (int, int*)`, где ключ – номер файлового дескриптора, а значение – массив открытых процессом файловых дескрипторов).

В итоге в результате апробации программной реализации способа на ряде примеров было установлено, что при обеспечении соответствия условий и результатов выполнения событий «низкоуровневого» описания модели и заданных ACSL-спецификациями предусловий и постусловий функций сгенерированного Си-кода, можно утверждать, что успешная верификация модели на Event-B гарантирует верификацию автоматически сгенерированного Си-кода с использованием ACSL-спецификаций.

Заключение

Таким образом, в настоящей статье рассмотрен способ для проверки корректности реализации непосредственно в программном коде ОСН Astra Linux математической МРОСЛ ДП-модели, описывающей механизмы управления доступом ОСН. Данный способ включает автоматическую генерацию Си-кода на основе формализованного «низкоуровневого» описания модели на языке формального метода Event-B (этап 1) и автоматическую генерацию ACSL-спецификаций на основе сгенерированного Си-кода (этап 2). Программная реализация этапа 1 практически завершена. В настоящее время ведутся работы на этапе 2. Применение описанного способа позволит минимизировать ручной труд специалистов в процессе верификации кода с помощью формального языка ACSL, предоставив сгенерированные верифицированные функции на языке программирования Си, реализующие в системном ПО механизмы защиты. Одним из перспективных направлений развития данного исследования является генерация Си-кода, реализующего ролевое управление доступом, и дальнейшее его использование в ядре будущих релизов ОСН Astra Linux.

Список литературы

1. Девянин П.Н., Хорошилов А.В., Тележников В.Ю. Формирование методологии разработки безопасного системного программного обеспечения на примере операционных систем // Труды Института системного программирования РАН, 2021. Т. 33, № 5. С. 25-40.
2. ГОСТ Р 59453.1-2021 «Защита информации. Формальная модель управления доступом. Часть 1. Общие положения». М.: Стандартинформ. 16 с.
3. ГОСТ Р 59453.2-2021 «Защита информации. Формальная модель управления доступом. Часть 2. Рекомендации по верификации формальной модели управления доступом». М.: Стандартинформ. 12 с.
4. Девянин П.Н. Модели безопасности компьютерных систем. Управление доступом и информационными потоками. Учебное пособие для вузов. 3-е изд., перераб. и доп. М.: Горячая линия – Телеком, 2020. 352 с.: ил.
5. Девянин П.Н., Тележников В.Ю., Третьяков С.В. Основы безопасности операционной системы Astra Linux Special Edition. Управление доступом. Учебное пособие / Под ред. чл.-корр. Академии криптографии России, доктора техн. наук, профессора П.Н. Девянина. – М.: Горячая линия – Телеком, 2022. – 148 с.: ил.
6. Abrial J.-R., Butler M., Hallerstede S. et al. Rodin: An Open Toolset for Modelling and Reasoning in Event-B // International Journal on Software Tools for Technology Transfer. 2010. Vol. 12, no. 6. P. 447–466.
7. Девянин П.Н., Леонова М.А. Приёмы по доработке описания модели управления доступом ОСН Astra Linux Special Edition на формализованном языке метода Event-B для обеспечения её

автоматизированной верификации с применением инструментов Rodin и ProB // Прикладная дискретная математика, 2021. № 52. С. 83-96.

8. Кокорин А.О., Тиевский С.Д., Девянин П.Н. Приемы дедуктивной верификации программного кода с использованием AstraVer Toolset // Прикладная дискретная математика. Приложение. 2022. № 15. С. 80-90.

9. ANSI/ISO C Specification Language. 2022. URL: <https://frama-c.com/html/acsl.html>.

10. Ефремов Д.В., Копач В.В., Корныхин Е.В., Кулямин В.В., Петренко А.К., Хорошилов А.В., Щепетков И.В. Мониторинг и тестирование модулей операционных систем на основе абстрактных моделей поведения системы // Труды Института системного программирования РАН. 2021, Т. 33, № 6. С. 15-26.

АСИМПТОТИКА РЕШЕНИЙ УРАВНЕНИЙ НА СЕТКАХ В ДИСКРЕТНОМ ВРЕМЕНИ

Михайлова С.О., Петухова Е.С.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ*

Аннотация

В работе на примере задачи об одномерном случайном блуждании развивается метод построения неосциллирующих ВКБ решений, описывающих процесс случайного блуждания. Уравнение на сетке продолжается на всю прямую и рассматривается полученное псевдодифференциальное уравнение. Для него строится фундаментальное решение задачи Коши, а также находится оценка разности между точным и асимптотическим ВКБ решением.

Введение

Целью данной работы является построение фундаментального решения задачи Коши для продолженного уравнения и нахождение точности асимптотического решения задачи о случайном блуждании по решетке. Асимптотическое решение строится с помощью неосциллирующего метода Вентцеля-Крамерса-Бриллюэна (далее ВКБ) в форме, пригодной для исследования псевдодифференциальных уравнений параболического типа.

В данной работе случайные блуждания характеризуются вероятностью попадания в точку в определенный момент времени. Для этого используется следующий прием. Пусть U_i (где i – номер узла) – функция, заданная на сетке, $U(x)$ – функция на прямой, такая что $U(ih) = U_i$ при всех значениях i (где h – шаг сетки). Тогда $U_{i+k} = U(ih + kh) = \exp\left(kh \frac{\partial}{\partial x}\right) U(x)|_{x=ih}$. Таким образом, любая линейная комбинация значений U_i может быть представлена в виде действия псевдодифференциального

оператора (линейная комбинация операторов сдвига) на функцию $U(x)$.

Этот подход использовался ранее для исследования гиперболических разностных уравнений [1], [2], [8]. При этом для построения $U(x)$ была использована интерполяционная формула, связанная с теоремой Котельникова. Эта формула основана на преобразовании Фурье, содержащем экспоненту с мнимым показателем. Такую же экспоненту содержит и представление -функции Дирака в виде преобразования Фурье [3], [6], [7]. Поэтому оба эти представления не подходят для построения неосциллирующих ВКБ решений в виде: $\varphi(x, t, h)e^{\frac{-S(x, t)}{h}}$, где показатель экспоненты – вещественная функция.

Операторное представление -функции Дирака получено давно и многократно использовалось. Замена интерполяционной функции Котельникова изложена в разделе «Связь между решениями псевдодифференциального и разностного уравнений».

Асимптотика вероятностей задачи о случайном блуждании актуальна тема. Так в [4] исследуется асимптотика поведения частицы при случайном блуждании с особыми временными паттернами. В [5] приведена асимптотика для случайных блужданий с непрерывным временем. Асимптотика случайных блужданий, которые можно описать, дифференциальными уравнениями дробного порядка исследуется в [6].

Случайные блуждания по целым числам с дискретным временем хорошо исследованы в рамках вероятности выхода за пределы области и первого попадания в точку [10], [11]. В данной работе задача рассматривает в терминах вероятности нахождения частицы в определенный момент времени, в определенной точке в пространстве.

Постановка задачи

Рассмотрим случайное блуждание частицы по одномерной решетке (дискретной подгруппе \mathbb{R}_1) с дискретным временем, где h – фиксированная длина -го направленного шага блуждания; а сам k -й направленный шаг блуждания $x_k = -h$, когда частица делает шаг влево, и $x_k = h$, когда частица делает шаг вправо. Обозначим за S_k координату частицы на k -м шаге блуждания. Очевидно, что $\sum_{i=0}^n x_i = S_k$. Обозначим через U_k^n вероятность в момент времени n нахождения частицы, выпущенной в нулевой момент времени из положения $S_0 = 0$, в положении S_k :

$$P\left(\sum_{i=0}^n x_i = S_k\right) = U_k^n. \quad (1)$$

Таким образом, простейшее случайное блуждание частицы, которая может на каждом шаге перейти в соседние состояния, можно переписать в силу формулы полной вероятности в виде:

$$U_k^{n+1} = \frac{\lambda + 1}{2} U_{k-1}^n - \frac{\lambda - 1}{2} U_{k+1}^n \quad (2)$$

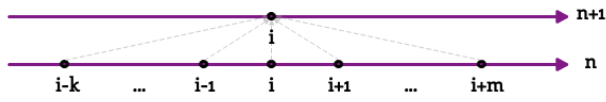


Рис. 1. Случайное блуждание по целочисленной решетке

Аналогично блуждание по одномерной решетке с конечным числом шагов (вправо и влево) с некоторой вероятностью (Рис. 1) описывается разностными уравнениями вида:

$$U_i^{n+1} = a_{i-k,i} U_{i-k}^n + \dots + a_{i+m,i} U_{i+m}^n \quad (3)$$

где $a_{j,i}$ - вероятность перехода из состояния j в состояние i , такая что $a_{j,i} \geq 0$, $\sum_j a_{j,i} = 1$; величины $a_{j,i}$, U_i^0 - заданы.

Для блуждания по дискретной подгруппе \mathbb{R}_n будет верно равенство:

$$U_{\vec{i}}^{n+1} = a_{\vec{i}-\vec{k},\vec{i}} U_{\vec{i}-\vec{k}}^n + \dots + a_{\vec{i}+\vec{m},\vec{i}} U_{\vec{i}+\vec{m}}^n \quad (4)$$

где $a_{\vec{j},\vec{i}}$ — вероятность перехода из состояния $\vec{j} = (j_1, \dots, j_n)$ в состояние $\vec{i} = (i_1, \dots, i_n)$, такая что

$a_{\vec{j},\vec{i}} \geq 0$, $\sum_{j_1} \dots \sum_{j_n} a_{\vec{j},\vec{i}} = 1$; величины $a_{\vec{j},\vec{i}}$, $U_{\vec{i}}^0$ заданы.

Далее мы будем рассматривать одномерную задачу о случайном блуждании на основе техники псевдодифференциальных уравнений. Мы заменяем функцию дискретного аргумента U_n^i функцией $U(x, t)$, где τ - длина шага по времени, а уравнение на решетке заменяется псевдодифференциальным уравнением:

$$e^{\tau \frac{\partial}{\partial t}} U = \left(\sum_{j=-k}^m a_j(x) e^{jh \frac{\partial}{\partial x}} \right) U, \quad (5)$$

где $e^{kh \frac{\partial}{\partial x}}(f) = f(x + kh)$ – оператор сдвига. Далее будем предполагать, что $\tau = h$.

Связь между решениями псевдодифференциального и разностного уравнений

Решение разностного уравнения для вероятностей можно найти при помощи аппарата фундаментальных решений. Фундаментальное решение задачи Коши для псевдодифференциального уравнения обозначим за $G(x, \xi, t)$. Решение задачи Коши с начальными данным $U^0(x)$ можно представить в виде:

$$U(x, \xi, t) = \int_{\mathbb{R}_1} G(x, \xi, t) U^0(\xi) d\xi \quad (6)$$

Введем функции $\Pi_j = \begin{cases} 0, & x < j - \frac{h}{2}, j + \frac{h}{2} \leq x; \\ 1, & j - \frac{h}{2} \leq x < j + \frac{h}{2}. \end{cases}$

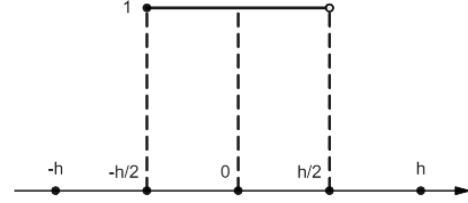


Рис.2. Функция Π_0

Заметим, что $\Pi_k = \Pi_0(x - kh)$ и $\sum_{k=-\infty}^{\infty} \Pi_k \equiv 1$ - образует разбиение единицы. При этом начальные данные равны $U^0(x) = \frac{1}{h} \Pi_0(x)$.

Дискретную вероятность U_j^n определим по формуле:

$$U_j^n = \int_{\mathbb{R}_1} U(x, nh) \Pi_j(x) dx = \quad (7)$$

$$= \frac{1}{h} \int_{\mathbb{R}_1} \int_{\mathbb{R}_1} \Pi_0(\xi) G(x, \xi, nh) \Pi_j(x) dx d\xi$$

При этом начальные данные для данной функции будут иметь вид:

$$U_j^0 = \frac{1}{h} \int_{\mathbb{R}_1} \Pi_j(x) \Pi_0(x) dx = \delta_{0,j} = \begin{cases} 1, & j = 0 \\ 0, & j \neq 0 \end{cases}.$$

Вероятность U_j^n будет обладать следующими свойствами.

Утверждение 1: обозначим $U^n(x) = U(x, nh)$ - решение уравнения (5) с постоянными коэффициентами и пусть $U^0(x) \geq 0$, $\int_{\mathbb{R}_1} U^0(x) dx = 1$. Тогда $U^n(x) \geq 0$, $\int_{\mathbb{R}_1} U^n(x) dx = 1$.

Утверждение 2:

$$\sum_{j=-\infty}^{\infty} U_j^n = 1 \quad (8)$$

в силу свойств разбиения единицы Π_j .

Для функции U_j^n , определенной в (7), верны основные свойства вероятности: (8) и неравенство $U_j^n \geq 0$.

Утверждение 3: для определенной в (7) функции U_j^n справедливо равенство $U_j^n = \sum_{k=-j_0}^{j_1} a_k U_{k+j}^n$.

В случае переменных коэффициентов $U^{n+1}(x, h) = \sum_{k=-j_0}^{j_1} a_j(x + jh) U^n(x + jh, h)$.

Утверждение 4: для определенной в (7) функции U_j^n и переменных коэффициентов $a_j(x)$ верно следующее $U_k^{n+1} = \sum_{j=-j_0}^{j_1} a_j((k+j)h)U_{k+j}^n + O(h^2)$.

Таким образом за счёт формулы (6) можно построить непрерывную функцию, отвечающую (5) и начальному условию. Переход от непрерывной функции к дискретной осуществляется по средствам формулы (7).

Оценка асимптотического решения

Асимптотическое решение задачи Коши для уравнения (5) можно построить в виде ВКБ:

$U_{as} = \varphi(x, t, h)e^{\frac{-S(x, t)}{h}}$, эта функция по построению удовлетворяет уравнению

$$e^{\frac{\partial}{\partial t}} U = \left(\sum_{j=-k}^m a_j(x) e^{jh \frac{\partial}{\partial x}} \right) U + O(h^2) e^{\frac{-S(x, t)}{h}}.$$

Функция $S(x, t)$ удовлетворяет уравнению Гамильтона-Якоби: $e^{-St} = \frac{\lambda+1}{2} e^{S_x} - \frac{\lambda-1}{2} e^{-S_x}$, полученному при подстановке асимптотического решения в исходное уравнение (5).

U_i^n есть сужение на сетку неосциллирующего ВКБ решения, поэтому в области $S \geq \delta > 0$ оценка $\max_i |U_i^n| \leq \max_i |U_i^0|$ и следующая из нее оценка для разности между точным и приближённым решением ничего не говорит о малости погрешности в области. В рассматриваемом случае следует получить оценку с весом равным $e^{\frac{S(x, t)}{h}}$. Для этого рассмотрим разность $W = e^{\frac{S(x, t)}{h}} (U - U_{as})$.

Теорема 1. Пусть функция $S(x, t)$ удовлетворяет уравнению Гамильтона-Якоби и ее вторые производные ограничены, тогда справедлива следующая оценка:

$$\max_i |W_i^{n+1}| \leq e^{\frac{1}{2} \int_0^t \max_x (S_{tt} - S_{xx}) dt'} (\max_i |W_i^0| + O(h)).$$

Таким образом если точное и асимптотическое решение совпадают в начальный момент времени, асимптотическое решение хорошо приближает точное, в случае, когда $e^{\frac{1}{2} \int_0^t \max_x (S_{tt} - S_{xx}) dt'}$ ограничено.

Заключение

Основной метод асимптотического анализа задачи о случайном блуждании основан на распространении дискретной задачи на все конфигурационное пространство, прямую в данном случае.

В этой работе была рассмотрена общая задача о случайных блужданиях, на примере двухшагового блуждания, и получена асимптотическая формула для вероятности нахождения частицы в фиксированный момент времени в фиксированной точке пространства. Асимптотическое решение задачи Коши построено неосциллирующим методом ВКБ.

Разность между точным и асимптотическим решениями по итогам исследования мала при предположении, что на интервале $[0, T]$, на котором построено асимптотическое решение, фазовая функция $S(x, t)$ — гладкая. Это значит, что асимптотическое решение можно принять за точное.

Результат этой работы применим для случаев, когда необходимо быстро получить приблизительное решение. Возможные применения этой оценки будут изучены в ходе дальнейших исследований.

Список литературы

1. Danilov V. G. A representation of the delta-function via creation operators and gaussian exponentials, and multiplicative fundamental solution asymptotics for some parabolic pseudodifferential-equations //Russian Journal of Mathematical Physics. – 1995. – V. 3. – №. 1. – P. 25-40.
2. Danilov V. G. Nonsmooth nonoscillating exponential-type asymptotics for linear parabolic PDE //SIAM Journal on Mathematical Analysis. – 2017. – V. 49. – №. 5. – P. 3550-3572.
3. Danilov V. G., Frolovitch S. M. Exact asymptotics of the density of the transition probability for discontinuous Markov processes //Mathematische Nachrichten. – 2000. – V. 215. – №. 1. – P. 55-90.
4. Magdziarz M., Szczotka W., Żebrowski P. Asymptotic behaviour of random walks with correlated temporal structure //Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. – 2013. – V. 469. – №. 2159. – P. 20130419.
5. Shlesinger M. F. Asymptotic solutions of continuous-time random walks //Journal of Statistical Physics. – 1974. – V. 10. – №. 5. – P. 421-434.
6. Umarov S., Steinberg S. Random walk models associated with distributed fractional order differential equations //Lecture Notes-Monograph Series. – 2006. – P. 117-127.
7. Маслов В. П., Назайкинский В. Е. О применении канонического оператора в термодинамике //Доклады Академии наук. – Федеральное государственное бюджетное учреждение «Российская академия наук», 2006. – Т. 411. – №. 1. – С. 12-18.
8. Маслов В. П. Асимптотические методы и теория возмущений. – «Наука», Главная редакция физико-математической литературы, 1988.
9. Маслов В. П., Федорюк М. В. Квазиклассическое приближение для уравнений квантовой механики. – Наука. Гл. ред. физ.-мат. лит., 1976. – С. 296.
10. Феллер В. Введение в теорию вероятностей и ее приложения. – Рипол Классик, 2013.
11. Ширяев А. Н. Вероятность. – МЦНМО, 2007.

РЕГУЛЯРИЗОВАННЫЙ ФИЛЬТР КАЛМАНА В ПРИЛОЖЕНИИ К ЧАНДЛЕРОВСКОМУ ДВИЖЕНИЮ ПОЛЮСА

Гаража И.А., Барина С.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В работе предлагается метод регуляризации фильтра Калмана (ФК), основанный на рассмотрении его как расширения рекуррентного метода наименьших квадратов (МНК). Во многих работах предпринимались попытки регуляризовать ФК, которые обычно сводились к добавлению стабилизирующего члена в матрицу усиления. Нашей целью являлось получение более общих математических выражений и их применение к задаче оценивания Чандлеровского движения полюса Земли и входного возбуждения для него. С учетом погрешностей наблюдений в разные эпохи, проанализировано движение полюса с 1840-х гг. Выявлено, что в 2017-2020-е гг. Чандлеровская компонента практически затухла.

Введение

Траектория полюса вращения Земли имеет выраженные годовую, вековую и чандлеровскую составляющие. Чандлеровское движение полюса (ЧДП) имеет период 433 суток. Оно было открыто С. Чандлером в 1891 г. За более чем столетие наблюдений выявлено следующее необычное поведение этой компоненты: его амплитуда колеблется в диапазоне от 100 до 300 миллисекунд дуги (от 3 до 10 метров) с характерными временами около 40 лет, в 1930-е годы амплитуда ЧДП резко уменьшалась, а фаза изменилась на 180° [11].

Для установления причин, вызывающих изменение амплитуды ЧДП, требуется уметь вычислять входное воздействие, вызывающее такие изменения, и сопоставить его с другими процессами, проходящими на Земле или в космосе. При этом, наблюдения проводились в разные эпохи с разной точностью, а задача нахождения входного воздействия динамической системы относится к категории некорректно поставленных (по Ж. Адамару) обратных задач.

В данной работе в разделе «Регуляризованный фильтр Калмана» представлена регуляризация ФК, аналогичная регуляризации рекуррентного МНК. В разделе «Методика оценки входного воздействия в модели движения полюса Земли» описано использование ФК для динамической системы движения полюса с регуляризацией по уравнению состояния. В разделе «Полученные результаты» дано решение задачи оценивания ЧДП и входного воздействия в чандлеровском диапазоне, представлены итоговые графики, сопровождаемые «Заключением».

Регуляризованный фильтр Калмана

Пусть дана дискретная линейная динамическая система:

$$\begin{aligned}x(j+1) &= A(j)x(j) + q(j), \\ z(j) &= C(j)x(j) + r(j),\end{aligned}$$

где $x(j) - n \times 1$ - вектор состояния,

$A(j) - n \times n$ матрица, связывающая текущее состояние с предыдущим,

$q(j) - n \times 1$ вектор белого шума с нулевым средним и корреляционной матрицей Q для нулевой задержки,

$z(j) - l \times 1$ вектор наблюдений,

$C(j) - l \times n$ матрица наблюдений,

$r(j) - l$ -мерный шум наблюдений (некоррелированный), с матрицей R для нулевой задержки.

С помощью ФК, используя уравнения состояния и наблюдений, оценивается вектор состояния динамической системы, при условии наличия сведений о ковариационных свойствах шумов наблюдений и входных воздействий [1-2]. В настоящее время алгоритм получил широкое распространение, существует множество его модификаций [3-6]. В курсе лекций Пантелеева В.Л. [7] и книге Губанова В.С. [11] дискретный фильтр Калмана описан как расширение рекуррентного метода наименьших квадратов (МНК).

Матрица усиления в ФК вычисляется по формуле

$$K = P_j^- C^T (C P_j^- C^T + R)^{-1},$$

где P_j^- - ковариационная матрица погрешности априорной оценки.

С вероятностной точки зрения коэффициент K есть не что иное, как матричная ковариация. По аналогии с рекуррентным МНК. на каждом шаге имеется оценка вектора состояния x^- и его ковариационная матрица P_x^- . Далее требуется скорректировать оценку на основе рассогласования нового наблюдения z . При этом строится уравнение линейной регрессии и оцениваются его параметры.

Алгоритм позволяет уточнить вектор состояния

$$x^+ = x^- + K(z - Cx^-),$$

по вновь поступившему наблюдению z с уточнением ковариационной матрицы

$$P_j^+ = P_j^- + K_j C_j P_j^{-1}.$$

Уравнение состояния используется для перехода на следующий шаг по времени

$$\begin{aligned}x_{j+1}^- &= A_j x_j^+ \\ P_{j+1}^- &= A_j P_j^+ A_j^T + Q_j.\end{aligned}$$

Рассматривая фильтр Калмана как расширение рекуррентного МНК, добавим Тихоновский регуляризатор к нормальной матрице системы, которая по-

является при корректировке решения с учетом наблюдений. Если нормальная матрица системы в методе МНК $C^T C$ плохо обусловлена, то она может быть скорректирована добавлением члена αI , где α – параметр регуляризации, I – единичная матрица соответствующей размерности:

$$P_N^- = \sigma_0^2 (C^T C + \alpha I)^{-1},$$

и апостериорной (по $N+1$ наблюдению)

$$P_{N+1}^+ = \sigma_0^2 (C^T C + c_{N+1}^T c_{N+1} + \alpha I)^{-1}.$$

Здесь σ_0^2 – дисперсия единицы веса, c_{N+1} – строка коэффициентов $N+1$ -го уравнения наблюдений. При этом выражение для матрицы усиления останется без изменений

$$K(N) = ((C^T C + \alpha I)^{-1} c_{N+1}^T (1 + c_{N+1} (C^T C + \alpha I)^{-1} c_{N+1}^T)^{-1}) = P_N^- c_{N+1}^T (c_{N+1} P_N^- c_{N+1}^T + \sigma_0^2)^{-1}.$$

В нём стоят регуляризованные ковариационные матрицы, использованные вместо обратных нормальных матриц. В рамках такого формализма уточненные оценки могут быть получены по не изменившимся формулам (см [7] гл. 11):

$$\hat{x}(N+1) = \hat{x}(N) + K(N)(z_{N+1} - c_{N+1} \hat{x}(N)),$$

$$P_{N+1}^+ = (E - K(N)c_{N+1})P_N^-.$$

Регуляризованный фильтр Калмана уже применялся к решению прикладных задач в разных областях [8-10]. Однако, регуляризация при этом делалась добавлением регуляризующего члена в формулу матрицы усиления в виде

$$K = P_j^- C^T (C P_j^- C^T + R + \alpha I)^{-1}.$$

На наш взгляд такая формула может быть получена лишь приближенно, в ходе отбрасывания некоторых членов в полученных нами более общих формулах. Однако написанная выше формула вполне применима, из нее видно, что ковариационную матрицу погрешностей наблюдений R можно использовать как стабилизатор.

Методика оценка входного воздействия в модели движения полюса Земли

Рассмотрим в качестве модели движения полюса линеаризованные уравнения Эйлера-Лиувилля [11]:

$$DI \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{pmatrix} -\beta & -\alpha \\ \alpha & -\beta \end{pmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \begin{pmatrix} \beta & \alpha \\ -\alpha & \beta \end{pmatrix} \begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix},$$

где α, β – действительная и мнимая составляющие комплексной чандлеровской частоты

$$\sigma_c = \alpha + i\beta = 2\pi f_c + i \frac{\pi f_c}{Q},$$

Q – добротность (примем $Q \approx 100$ [14]), $f_c = 1/433$ сут⁻¹ – действительная чандлеровская частота, $D = \frac{d}{dt}$ – оператор дифференцирования, χ_1, χ_2 – компоненты входного возбуждения, i – мнимая единица, I – единичная матрица $[2 \times 2]$ [12].

Для оценки входного воздействия по известным координатам полюса Земли при постоянных параметрах уравнений связь χ с p можно переписать через обратную передаточную функцию $W^{-1}(D)$:

$$\begin{bmatrix} \chi_1 \\ \chi_2 \end{bmatrix} = \begin{pmatrix} \beta & \alpha \\ -\alpha & \beta \end{pmatrix}^{-1} \left(DI - \begin{pmatrix} -\beta & -\alpha \\ \alpha & -\beta \end{pmatrix} \right) \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = W^{-1}(D) \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}.$$

Обратный оператор $W^{-1}(D)$, содержащий дифференцирование p , усиливает любые шумы наблюдений вне резонансной частоты f_c . Поэтому необходима либо регуляризация, либо дополнительная фильтрация наблюдений p , прежде чем выделять из них χ . Встает необходимость регуляризации задачи не по уравнению наблюдений, а по уравнению состояний.

Для решения проблемы возможно применение нескольких подходов. В рамках корректирующей фильтрации В.Л. Пантелеева [13] частотный фильтр $W_{\text{corr}}(f)$ может быть применен к p до вычисления входного воздействия χ , что позволяет, дополнительно отфильтровав наблюдения вблизи чандлеровского диапазона, предотвратить просачивание шумов и их усиление обратным оператором $W^{-1}(f)$.

Корректирующий фильтр можно сформировать также на основе Тихоновской регуляризации:

$$W_{\text{corr}}(f)W^{-1}(f) = \frac{W^*(f)}{W^*(f)W(f) + \alpha I},$$

где $*$ обозначает сопряжение, оператор D заменен на $i2\pi f$. С помощью параметра регуляризации происходит отсеечение составляющих, близких к или принадлежащих к ядру оператора W , тем самым сужается частотная область решения χ .

Учитывая различие в тысячу раз точности наблюдений в разные эпохи (в XIX веке погрешность была 0.1, а в XXI веке – 0.0001 сек дуги), целесообразно несколько регуляризовать матрицу $R(t)$, чтобы веса наблюдений, берущиеся обратно пропорциональными квадрату погрешности, не отличались в миллионы раз. Мы сделали это, взяв их в степени -1.5.

Полученные результаты

На рис. 1, а представлен результат выделения ЧДП фильтром Калмана, а на рис. 1, б показано входное возбуждение для него, обработанное фильтром Пантелеева, в сравнении с суммарным угловым моментом атмосферы и океана (ААМ+ОАМ) по данным реанализов GFZ. Представлены компоненты по x координате (по y они аналогичны).

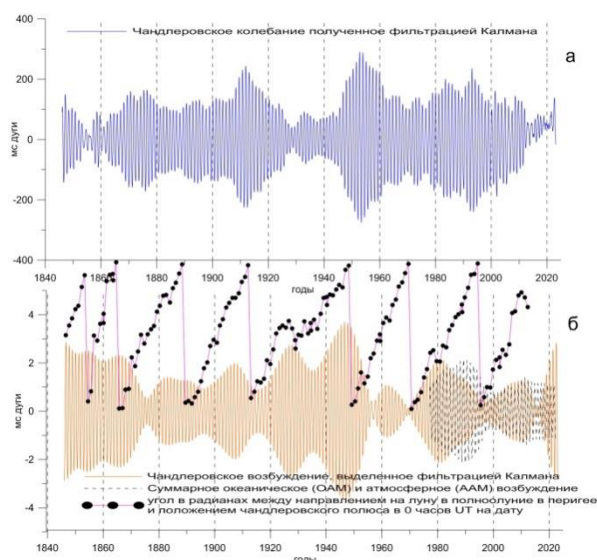


Рис. 1. а) Выделенное чандлеровское колебание p_1 .
б) Чандлеровское возбуждение χ_1 в сравнении с суммарным угловым моментом атмосферы и океана ААМ+ОАМ GFZ. Приведен также график изменения угла между направлением на Чандлеровский полюс и перигей лунной орбиты в большие полнолуния

Это стандартный подход сравнения полученных из движения полюса (геодезического) и по метеорологическим данным (геофизического) возбуждений говорит о том, что влияние океана и атмосферы не вполне согласуются с возбуждением χ_1 , а значит и не вполне объясняют резкое уменьшение амплитуды ЧДП в 2017-2020 гг.

В то же время, в последних работах Л.В. Зотов установил, что причиной модуляций как чандлеровского колебания, так и скорости вращения Земли с 22-летними и 90-летними квази-периодами может являться смена сизигиев (полнолуний и новолуний) в перигее, особенно сильно сказывающихся на всех составляющих лунно-солнечных приливов. Направление на перигей лунной орбиты, если сопоставлять его в дни этих событий (в одно и то же время, например, в Гринвичскую полночь) с направлением смещения полюса, будет описывать полный круг (цикл) за ~22 года, о чем свидетельствует соответствующая кривая изменения этого угла, представленная на рис 16. Изменение фазы чандлеровского колебания на 180° в 1930-е гг., наметившееся и в 2020-е, может быть результатом того, что большие сизигии дрейфуют на фоне 18.6-летнего цикла прецессии за 90 лет. При этом по истечении 22- и 90-летнего циклов полнолуния и новолуния в перигее чередуются [15].

Заключение

В работе рассмотрен новый подход к регуляризации фильтра Калмана. Этот метод отличается от имеющихся в литературе [8-10] тем, что он основан на рекуррентном МНК и включает стабилизатор как в кросс-ковариационную, так и в обратную дисперсионную матрицы в формуле усиления K .

Показано, что при решении задачи оценивания ЧДП и входного возбуждения для него требуется регуляризация по уравнению состояний, которая может выполняться с использованием корректирующего фильтра Пантелеева.

Получена траектория ЧДП и возбуждение для него, выполнено сравнение с суммарным угловым моментом океана и атмосферы. Дан краткий экскурс в новую концепцию Л.В. Зотова, способную пролить свет на причины модуляций в чандлеровском колебании полюса Земли.

Список литературы

1. Kalman R.E. A new approach to linear filtering and prediction problems. Transactions ASME-J Basic Eng., 82, Series D, 35–45, 1960
2. Калман Р., Фолб П., Арбиб М. Очерки по математической теории систем. М., УРСС, 2004
3. Фомин В.Н. Оптимальная и адаптивная фильтрация. СПбГУ, 2003
4. Ширяев А.Н. Вероятность 1,2. М., МЦМНО, 2007
5. Арато М. Линейные стохастические системы с постоянными коэффициентами. М., Наука, 1989
6. Зотов Л.В. Теория фильтрации и обработка временных рядов. М., МГУ, 2010
7. Пантелеев В.Л. Наблюдение и управление динамическими объектами, курс лекций, ГАИШ МГУ, 2001
8. Li YM, Gu QM, Han SH, Gu YW, Tikhonov Regularized Kalman Filter and its Applications in Autonomous Orbit Determination of BDS., WSEAS Transactions on Mathematics archive, Vol 16, 2017, pp. 187-196
9. Трубин В. А., Батищев В. Регуляризованный фильтр Калмана. Патент 1317652 СССР, Опубл. в Б. И. N 7, 1988
10. Asgharpour M., Foodeh R., Daliri M.R. Regularized Kalman filter for brain-computer interfaces using local field potential signals, J Neurosciences Methods. N 350, 2021, 109022
11. Губанов В.С. Обобщенный метод наименьших квадратов. Теория и применение в астрометрии. СПб. Наука, 1997
12. Зотов Л.В. Климатические процессы и вращение Земли. М. МИЭМ, 2022, ISBN:978-5-600-03155-5
13. Zotov L., Panteleev V.L. Filtering and inverse problems solving, in Computational Methods for Applied Inverse Problems Edited by Y. F. Wang, A. G. Yagola and C. C. Yang, De Gruyter & Higher Education Press, June 2012, pp. 169-194
14. Nastula J., and R. Gross, Chandler wobble parameters from SLR and GRACE. J. Geophys. Res. Solid Earth, 120, 4474–4483. 2015 doi: 10.1002/2014JB011825.
15. Zotov L., Ch. Bizouard, N. Sidorenkov, C.K. Shum, S. Denisenko Regularized Kalman filter in application to the Chandler wobble of the pole, ICCC workshop 2023, talk S05C11, online, 29/03/2023.

ПОСТРОЕНИЕ И ОБОСНОВАНИЕ АСИМПТОТИКИ ФУНДАМЕНТАЛЬНОГО РЕШЕНИЯ ДЛЯ ВЫРОЖДАЮЩЕГОСЯ ОДНОМЕРНОГО ПАРАБОЛИЧЕСКОГО УРАВНЕНИЯ С МАЛЫМ ПАРАМЕТРОМ

Рахель М.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматривается задача о построении и обосновании асимптотического (приближенного) фундаментального решения дифференциального уравнения в частных производных второго порядка параболического типа с малым параметром и вырождением при старшей производной. Построение асимптотики основано на операторном представлении дельта-функции Дирака и методе ВКБ. Обоснование происходит путем построения формального ряда для асимптотики, оценки его общего члена и доказательства сходимости.

Введение

Построение асимптотических решений - один из наиболее мощных методов конструктивного исследования дифференциальных уравнений в частных производных. В теории гиперболических уравнений такие методы хорошо известны, активно применяются и развиваются. Примерами являются интегральный оператор Фурье и канонический оператор Маслова. Такие теории используют объекты и понятия симплектической геометрии - лагранжевы многообразия и симплектические формы в фазовом пространстве. Для параболических уравнений подобные теории развиты слабее, и, хотя известен вариант теории канонического оператора Маслова [5], само направление содержит много открытых проблем.

Задача состоит в построении и обосновании асимптотики фундаментального решения задачи Коши для вырождающегося параболического уравнения с малым параметром, то есть отыскании функции $G_\varepsilon(x, \xi, t)$, удовлетворяющей задаче:

$$\begin{cases} \frac{\partial G}{\partial t} - \varepsilon \frac{\partial}{\partial x} \left(x^2 a^2(x) \frac{\partial G}{\partial x} \right) + b(x) \frac{\partial G}{\partial x} = 0, \\ G|_{t=0} = \delta(x - \xi), \end{cases} \quad (1)$$

где $a(x)$ - гладкая, положительная и ограниченная сверху вместе с производными функция, а условия на функцию $b(x)$ будут приведены в соответствующем разделе.

Можно отметить, что такие уравнения часто встречаются в теории случайных процессов, например, частным случаем уравнения (1) является уравнение, описывающее процесс Орнштейна-Уленбека [3].

Работа состоит из следующих разделов. В первом разделе напомним схему построения асимптоти-

ки фундаментального решения с помощью операторного представления δ -функции. В следующем разделе рассмотрен случай $b(x) \equiv 0$, для которого выписано главное слагаемое асимптотики фундаментального решения вместе с его свойствами. Также выписан формальный ряд для асимптотики, представлена оценка общего члена такого ряда, а также доказана его сходимость. Наконец, в последнем разделе будет сказано несколько слов о случае $b(x) \neq 0$, для которого будет выписан и просуммирован формальный ряд для асимптотики фундаментального решения.

Работа основана на совместной с Даниловым В. Г. статье [3].

Построение асимптотики

Ключевую роль в построении асимптотики фундаментального решения играет операторное представление дельта-функции Дирака в виде действия функции от оператора на гауссову экспоненту:

$$\delta(x) = \frac{1}{\sqrt{2\pi\varepsilon}} \lim_{\beta \rightarrow 1-0} e^{-\frac{\beta(x-a^+)^2}{2\varepsilon}} e^{-\frac{\zeta^2}{\varepsilon}}|_{\zeta=0}, \quad (2)$$

где a^+ - оператор рождения, например, $a^+ = \zeta - \frac{\hbar}{2} \frac{\partial}{\partial \zeta}$. Данное утверждение вместе с доказательством было приведено в [1].

Рассмотрим задачу для нахождения функции Грина параболического уравнения:

$$-\left[\varepsilon \frac{\partial}{\partial t} + H\left(x, -\varepsilon \frac{\partial}{\partial x}\right)\right] G = 0, \quad G|_{t=0} = \delta(x - \xi), \quad (3)$$

где $x \in \mathbb{R}, t \in [0, T]$. Определим символ фундаментального решения как решение задачи Коши:

$$-\left[\varepsilon \frac{\partial}{\partial t} + H\left(x, -\varepsilon \frac{\partial}{\partial x}\right)\right] V = 0, \quad V|_{t=0} = e^{-\frac{(x-\xi-y)^2}{2\varepsilon}}. \quad (4)$$

Тогда из (2) следует, что если мы знаем решение задачи (4) - функцию $V(x, \xi + y, t)$, то:

$$G(x, \xi, t) = \frac{1}{\sqrt{2\pi\varepsilon}} \lim_{\beta \rightarrow 1-0} V(x, \xi + a^+, t) e^{-\frac{\zeta^2}{\varepsilon}}|_{\zeta=0}. \quad (5)$$

Символ предлагается искать с помощью неосциллирующего метода ВКБ, то есть искать асимптотические приближения вида:

$$V_N = e^{-\frac{\Phi}{\varepsilon}} (\varphi_0 + \varepsilon \varphi_1 + \dots \varepsilon^N \varphi_N + O(\varepsilon^{N+1})). \quad (6)$$

Действуя по стандартной схеме, то есть подставив выражение для символа (6) в задачу (5), получим уравнение Гамильтона-Якоби для фазы (показателя экспоненты в выражении (6)):

$$\frac{\partial \Phi}{\partial t} + H\left(x, \frac{\partial \Phi}{\partial x}\right) = 0, \quad \Phi|_{t=0} = \beta \frac{(x-\xi-y)^2}{2} \quad (7)$$

и уравнения переноса для каждой из функций $\varphi_1, \varphi_2, \dots$:

$$\begin{cases} \Pi\varphi_0 = 0, \quad \varphi_0|_{t=0} = 1; \\ \Pi\varphi_k = F_k, \quad \varphi_k|_{t=0} = 0, \quad k \geq 1, \end{cases} \quad (8)$$

где оператор переноса Π и F_k будут выписаны ниже для каждого случая в отдельности.

Далее, используя операторное представление дельта-функции, можно вычислить действие оператора рождения в (5) и получить формулу для асимптотики фундаментального решения. В частности, формула для главного слагаемого имеет вид:

$$G_0(x, \xi, t) = \frac{1}{\sqrt{2\pi\varepsilon}} \lim_{\beta \rightarrow 1-0} \frac{e^{\frac{S(x, \xi, t)}{\varepsilon}} e^{\frac{1}{2} \int_0^t H_{xp} dt'}}{\sqrt{1 - \Phi''_{yy}}} \frac{1}{\sqrt{J_0}}, \quad (9)$$

где $S(x, \xi, t) = \hat{y}^2/2 - \Phi(x, \xi + \hat{y}, t)$ – решение уравнения Гамильтона Якоби (7), \hat{y} находится как решение неявного уравнения $\hat{y} = \Phi'_y(x, \xi + \hat{y}, t)$, а якобиан $J_0 = \frac{\partial x(x_0(x, \xi, t), \xi, t)}{\partial x_0}$.

Существует также другой способ перехода от уравнения для символа к фундаментальному решению, а именно к определению действия через интеграл от лагранжиана вдоль траекторий соответствующей системы Гамильтона. В [2] была представлена следующая коммутативная диаграмма:

$$\begin{array}{ccc} \Phi & \xrightarrow{L_{x \rightarrow p}} & \tilde{\Phi} \\ \downarrow y^+ & & \downarrow y^- \\ S & \xleftarrow{L_{p \rightarrow x}} & \tilde{S} \end{array}$$

где $L_{x \rightarrow p}, L_{p \rightarrow x}$ – преобразования Лежандра и

$$\begin{cases} y^- = \Phi(x, \xi + y, t) \rightarrow \Phi(x, \xi + \hat{y}, t) - \frac{\hat{y}^2}{2}, \\ \hat{y} = \frac{\partial \Phi}{\partial y}(x, \xi + \hat{y}, t); \\ y^+ = \tilde{\Phi}(p, \xi + y, t) \rightarrow \tilde{\Phi}(p, \xi + \tilde{y}, t) + \frac{\tilde{y}^2}{2} \\ \tilde{y} = -\frac{\partial \tilde{\Phi}}{\partial y}(p, \xi + \tilde{y}, t). \end{cases} \quad (10)$$

Смысл диаграммы состоит в том, что, используя левый (короткий) или правый (длинный) путь, мы получим одну и ту же функцию – показатель экспоненты в формуле (9).

Из свойств уравнения Гамильтона-Якоби (7) следует, что функция $\tilde{\Phi}$ будет решением уравнения $\frac{\partial \tilde{\Phi}}{\partial t} - H\left(\frac{\partial \tilde{\Phi}}{\partial p}, x\right) = 0$. Функция \tilde{S} также удовлетворяет этому уравнению, а различие будет заключаться только в функциях, отвечающих за начальные данные.

Начальные данные в задаче для $\tilde{\Phi}$ определяются соотношением $\tilde{\Phi}|_{t=0} = p^2/2 + p(\xi + y)$, а чит $S|_{t=0} = p\xi$.

Вырождающееся одномерное уравнение

Пусть $b(x) \equiv 0$. Характеристическая система уравнения Гамильтона-Якоби имеет вид:

$$\begin{cases} \dot{x} = 2x^2 a^2(x)p, \quad x|_{t=0} = \xi, \\ \dot{p} = -(x^2 a^2(x))'p, \quad p|_{t=0} = p_0. \end{cases} \quad (11)$$

Ее можно проинтегрировать, найти необходимую обратную проекцию и действие:

$$S(x, \xi, t) = \frac{1}{4t} \left(\int_{\xi}^x \frac{dz}{za(z)} \right)^2. \quad (12)$$

Решение первого уравнения переноса также может быть выписано явно:

$$\varphi_0(x, \xi, t) = \frac{1}{\sqrt{2\xi a(\xi)xa(x)t}}. \quad (13)$$

Главное слагаемое асимптотики фундаментального решения тогда будет иметь вид:

$$G_0(x, \xi, t) = \frac{1}{\sqrt{2\pi\varepsilon}} \frac{e^{-\frac{1}{4t\varepsilon} \left(\int_{\xi}^x \frac{dz}{za(z)} \right)^2}}{\sqrt{2\xi a(\xi)xa(x)t}}. \quad (15)$$

Такая функция G_0 будет решением задачи:

$$LG_0 = \left[\frac{\partial}{\partial t} - \varepsilon H \right] G_0 = -\varepsilon F_0, \quad G_0|_{t=0} = \delta(x - \xi), \quad (16)$$

где $H = \frac{\partial}{\partial x} \left(x^2 a^2(x) \frac{\partial}{\partial x} \right)$, $F_0 = G_0 F$.

Равенство (16) позволяет с помощью формулы Дюамеля написать формальный ряд для фундаментального решения в виде:

$$G_{as} = \sum_{k=0}^{+\infty} \varepsilon^k G_k(x, \xi, t), \quad (17)$$

где:

$$\begin{aligned} G_k(x, \xi, t) &= \\ &= \int_0^t \int_{-\infty}^{+\infty} G_0(x, y, t - \tau) L F_{k-1}(y, \xi, \tau) dy d\tau, \end{aligned} \quad (18)$$

а F_{k-1} – правая часть уравнения $L(G_0 + \varepsilon G_1 + \dots + \varepsilon^{k-1} G_{k-1}) = -\varepsilon^k F_{k-1}$.

Потребовав ограниченность функций $a(x), xa'(x), x^2 a''(x)$, можно доказать, что $F = \frac{\partial}{\partial x} \left(x^2 a^2(x) \frac{\partial \varphi_0}{\partial x} \right) / \varphi_0$ – равномерно ограниченная по x функция. Доказав также некоторые свойства свертки в (18), можно написать оценку для общего члена ряда (17). Она имеет вид:

$$G_k(x, \xi, t) \leq G_0(x, \xi, t) \frac{M^{2k} e^{2kM\varepsilon T} T^k}{k!}, \quad (19)$$

где $M = \max|F|$. Тогда для всего ряда верно:

$$G_{as} \leq G_0 \sum_{k=0}^{+\infty} \varepsilon^k \frac{M^{2k} e^{2kM\varepsilon T} T^k}{k!} = G_0 e^C, \quad (20)$$

где $C = \varepsilon M^2 T e^{2M\varepsilon T}$.

В итоге можно сформулировать следующее утверждение.

Теорема. Фундаментальное решение исходной задачи имеет вид:

$$G = \sum_{k=0}^{+\infty} \varepsilon^k G_k, \quad (21)$$

где ряд сходится равномерно при $t \in [0, T]$.

Вырождающееся уравнение со сносом

Пусть для простоты $a(x) \equiv 1$. Рассмотрим уравнение с гамильтонианом:

$$H(x, p) = x^2 p^2 + bxp. \quad (22)$$

Характеристическая система, соответствующая такому гамильтониану, также интегрируется. В этом случае:

$$S(x, \xi, t) = \frac{(\ln \frac{x}{\xi} - bt)^2}{4t}. \quad (23)$$

В решении уравнения переноса появляется множитель, отвечающий за снос:

$$\varphi_0(x, \xi, t) = \frac{e^{bt/2}}{\sqrt{2\xi xt}}. \quad (24)$$

Тогда главное слагаемое имеет вид:

$$G_0(x, \xi, t) = \frac{1}{\sqrt{2\pi\varepsilon}} \frac{e^{bt/2}}{\sqrt{2\xi xt}} e^{-\frac{(\ln \frac{x}{\xi} - bt)^2}{4t\varepsilon}}. \quad (25)$$

Также здесь могут быть найдены решения остальных уравнений переноса:

$$\varphi_k(x, \xi, t) = \frac{(-t/4)^k}{k!} \varphi_0(x, \xi, t), k \geq 1, \quad (26)$$

а значит точная асимптотика фундаментального решения будет иметь вид:

$$G(x, \xi, t) = G_0(x, \xi, t) \sum_{k=0}^{+\infty} \frac{(-\varepsilon t/4)^k}{k!} = G_0(x, \xi, t) e^{-\varepsilon t/4}. \quad (27)$$

Как было сказано выше, здесь приведены рассуждения для $a(x) \equiv 1$. Однако повторив те же выкладки, что и в предыдущем разделе, можно аналогичным образом построить формальный ряд для асимптотики и доказать его сходимости и для случая $a(x) \neq 1$. Различие будет заключаться только в множителе, отвечающем за снос.

Заключение

В работе была построена и обоснована точная асимптотика фундаментального решения вырождающегося параболического уравнения без сноса общего вида, а также рассмотрено уравнение со сносом в простейшем случае.

Так как для обоснования асимптотики в общем случае необходимо требовать лишь ограниченности функции $a(x)$ вместе с двумя ее первыми производными, данный подход можно без проблем перенести на более общий случай, например, большей размерности, что и предполагается сделать в дальнейшем.

Как было отмечено во введении, схожие задачи возникают также в теории вероятностей, а значит с помощью данного подхода можно строить решения

уравнений, описывающих различные случайные процессы и являющихся частными случаями уравнения (1).

Список литературы

1. Danilov V.G. A Representation of the Delta Function via Creation Operators and Gaussian Exponentials and Multiplicative Fundamental Solution Asymptotics for Some Parabolic Pseudodifferential Equations // Russ. J. Math. Phys. – 1995. – Vol. 3:1. – P. 25-40.
2. Danilov V.G., Frolovitch S.M. Exact Asymptotics of the Density of the Transition Probability for Discontinuous Markov Processes // Math. Nachr. – 2020. – Vol. 215. – P. 55-90.
3. Danilov V.G., Rakhel M.A. Development of Maslov's Approach to the Construction of Nonoscillating WKB-Type Solutions // Russian Journal of Mathematical Physics. – 2021. – Vol. 28:2. – P. 179-187.
4. Feller W. An Introduction to Probability Theory and Its Applications // John Wiley & Sons, Inc.; 2nd edition, 1971. – 751 p.
5. Maslov V.P., Nazaikinskii V.E. Tunnel canonical operator in thermodynamics // Funct. Anal. Appl. – 2006. – Vol. 40. – P. 173-187.

СВЯЗАННЫЕ СОСТОЯНИЯ В МОДЕЛИ ХАБДАРДА НА ЛЕСТНИЧНЫХ РЕШЕТКАХ

Пиле Я.Э.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В работе с помощью моделирования методом DMRG показан факт формирования связанных состояний трех фермионов (тримеров) в модели Хаббарда на лестничных решетках с тремя, четырьмя и пятью направляющими в случае неравных амплитуд туннелирования для носителей заряда с различным направлением спина. Результаты могут быть использованы для объяснения экспериментов в сверххолодных смесях атомных газов типа ${}^6\text{Li}$ - ${}^{40}\text{K}$.

Введение

Последние экспериментальные исследования сверххолодных двухкомпонентных поляризованных ферми-газов в низких размерностях [1-3] вызывают большой интерес, мотивированный поиском экзотических сверхтекучих фаз. Подобные фазы можно наблюдать в таких системах благодаря сильным корреляциям и квантовым флуктуациям. Особый интерес представляет состояние Фульде-Феррелла-Ларкина-Овчинникова (FFLO) [4,5] с ненулевым импульсом центра масс куперовских пар, в отличие от обычного режима Бардина-Купера-Шриффера. Теоретически было показано [6], что для одномерного

ферми-газа вся частично поляризованная фаза аналогична сверхпроводящему FFLO-состоянию и профиль плотности одномерного газа (который можно измерить напрямую) может служить доказательством этого [7]. Одним из наиболее широко используемых методов численного анализа таких систем остается приближение среднего поля [8]. Несмотря на некоторые плодотворные результаты в качественном объяснении экспериментальных данных, существуют значительные различия между решениями, полученными в приближении среднего поля и точным решением. В частности – фазовые диаграммы в моделях среднего поля [8,9] имеют две трикритические точки даже при нулевой амплитуде туннелирования (одномерный случай), в то время как точное решение с помощью анзаца Бете предсказывает только одну мультикритическую точку [7]. Одним из способов преодоления ограничений подходов среднего поля является использование методов численных расчетов высокой точности, например ренормализационной группы матрицы плотности (DMRG), который может быть использован для вычисления энергии и построения фазовых диаграмм для фермионов на лестничных решетках [11,12]. Существенной особенностью экспериментальных результатов для фазовых диаграмм является наличие одной мультикритической точки, в отличие от результатов, получаемых в приближении среднего поля. В работе [13] показано, что существование нескольких мультикритических точек связано с наличием связанных состояний трех фермионов (тримеров). В данной работе изучается наличие тримеров на лестничных решетках в модели Хаббарда в зависимости от различных параметров системы.

Модель

В качестве модели используется модель Хаббарда с притягивающим взаимодействием на лестничной решетке ширины W и длины L .

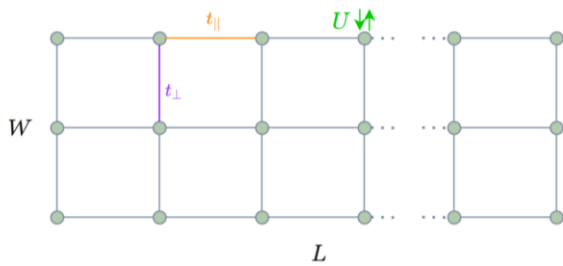


Рис. 1. Схематическое изображение лестничной решетки

Гамильтониан имеет вид:

$$\begin{aligned} \hat{H} = & -t_{\parallel,\sigma} \sum_{j=1,2} \sum_{i,\sigma} (c_{i,j,\sigma}^\dagger c_{i+1,j,\sigma} + h.c.) \\ & - t_{\perp,\sigma} \sum_{i,\sigma} (c_{i,j=1,\sigma}^\dagger c_{i,j=2,\sigma} + h.c.) \\ & + U \sum_{j=1,2} \sum_i n_{i,j,\uparrow} n_{i,j,\downarrow} \end{aligned}$$

Здесь $c_{i,j,\sigma}^\dagger/c_{i,j,\sigma}$ – операторы рождения-уничтожения электрона на узле с индексом i в продольном направлении, индексом j в поперечном направлении и спином σ , t_{\parallel} – амплитуда туннелирования в продольном направлении, t_{\perp} амплитуда туннелирования в поперечном направлении, U – потенциал контактного взаимодействия электронов, а $h.c.$ означает эрмитово сопряжение. В расчетах принято $t_{\parallel} = t_{\perp}$, а взаимодействие притягивающее ($U < 0$). Расчет энергии связи тримеров производился методом DMRG (Density Matrix Renormalization Group) – вариационным методом, который применяется для вычисления основного и низколежащих возбужденных состояний в квазиодномерных системах с высокой точностью [10]. Использована реализация метода в пакете iTensor (Algorithms and Libraries for Physics Simulations) и его интерфейс для языка C++ [14].

Расчеты

Для различных длин и конфигураций лестничных решеток посчитана энергия связи тримера. Расчет энергии основного состояния производится при фиксированном количестве электронов со спинами вверх и вниз. Если обозначить энергию системы с N_{\uparrow} электронами со спином вверх и N_{\downarrow} электронами со спином вниз за $E(N_{\uparrow}, N_{\downarrow})$, то энергия связи тримера будет иметь вид:

$$E_{trimer} = E(2,1) - E(1,1) - E(1,0).$$

Изменяемыми параметрами для рассматриваемых систем являлись длина лестницы L , характер граничных условий в поперечном направлении (открытые или периодические) и соотношение амплитуд туннелирования для электронов с различными направлениями спина. Для различных конфигураций была посчитана зависимость энергии связи тримеров от длины решетки L от 50 до 250.

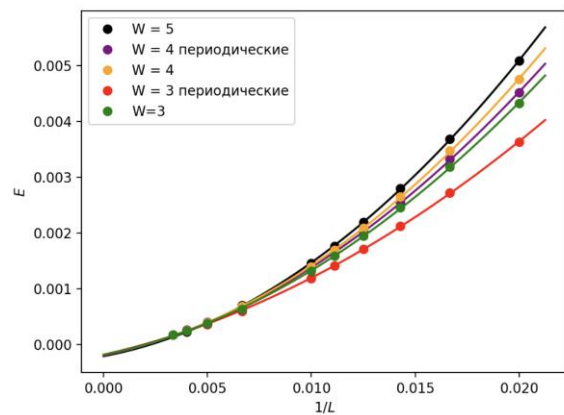


Рис. 2. Зависимость энергии связи тримеров от $1/L$ на лестницах с 3-5 направляющими с различными граничными условиями в поперечном направлении, $t_{\parallel,\sigma} = 1$, $t_{\perp,\sigma} = 1$, $U = -7t_{\parallel}$

На рис.2 приведена зависимость $E_{trimer}(1/L)$, точки – результат расчетов DMRG, а сплошные ли-

нии – результат квадратичной аппроксимации зависимости к пределу $L \rightarrow \infty$, реализованной на языке python с помощью метода наименьших квадратов из библиотеки numpy.

В пределе бесконечной длины значения функции (порядка 10^{-4}) не значимо отличаются от нуля (ниже уровня ошибки аппроксимации коэффициентов), что позволяет утверждать, что ненулевое значение энергии тримеров является результатом решеточных поправок из-за конечной длины.

Кроме того, интересным фактом является то, что поправки, связанные с конечной длиной, оказываются меньше в случае периодических граничных условий в поперечном направлении, что особенно хорошо видно по графикам для $W = 3$.

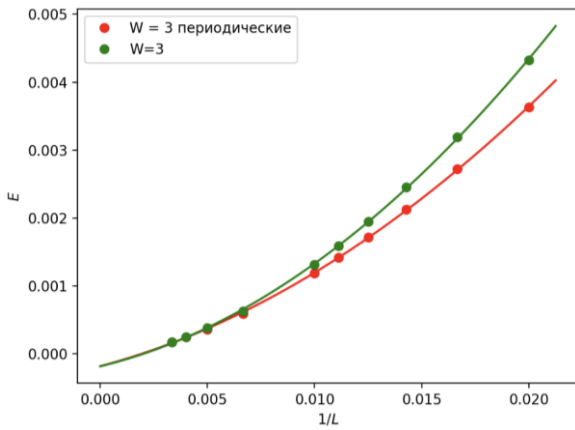


Рис. 3. Зависимость энергии связи тримеров от $1/L$ на лестницах с 3 направляющими с открытыми и периодическими граничными условиями в поперечном направлении, $t_{\parallel,\sigma} = 1$, $t_{\perp,\sigma} = 1$, $U = -7t_{\parallel}$

При переходе к амплитудам туннелирования, зависящим от направления спина, возникает ненулевая энергия связи тримеров.

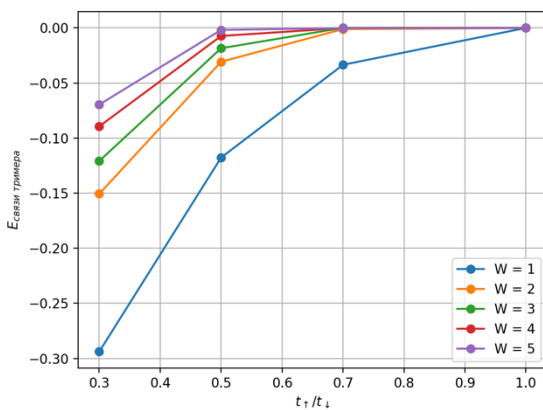


Рис. 4. Зависимость энергии связи тримеров от $t_{\uparrow}/t_{\downarrow}$ на лестницах с 1-5 направляющими с открытыми граничными условиями в поперечном направлении, $U = -7t_{\parallel}$

$$t_{\downarrow} = t_{\parallel,\downarrow} = t_{\perp,\downarrow} = 1 \text{ и } t_{\uparrow} = t_{\parallel,\uparrow} = t_{\perp,\uparrow}.$$

На рис.4 приведена зависимость $E_{\text{тример}}(t_{\uparrow}/t_{\downarrow})$, точки – результат квадратичной аппроксимации за-

висимости $E_{\text{тример}}(1/L)$, аналогичной рис.2. Видно, что чем меньше отношение амплитуд туннелирования, тем выше энергия связанных состояний (тримеров), кроме того, чем дальше система находится от строго одномерной конфигурации, тем больший дисбаланс необходим для появления связанных состояний.

Заключение

Получена зависимость энергии связи тримеров на лестницах с 1-5 направляющими в модели Хаббарда от длины лестницы, характера граничных условий в поперечном направлении и отношения амплитуд туннелирования электронов с различным направлением спина, для лестниц ширины больше 2 зависимость получена впервые. Экстраполяцией зависимости $E_{\text{тример}}(1/L)$ показано, что вычисленная энергия тримеров в пределе бесконечной длины имеет масштаб (порядка 10^{-4}) меньший, чем ошибка аппроксимации для случая спин-независимых амплитуд туннелирования и становится значимой (порядка 10^{-1}) при $t_{\uparrow}/t_{\downarrow} \sim 0.3$. Кроме того, показано, что решеточные поправки на конечную длину имеют меньший масштаб в случае периодических граничных условий в поперечном направлении. Известно также, что наличие тримеров в 1D разрушают дальний порядок в сверхпроводящей фазе [13], в качестве направления дальнейшей работы планируется рассмотреть, имеет ли место разрушение дальнего порядка в сверхпроводящей фазе на лестничных решетках, а также проверить наличие одной/нескольких мультикритических точек на фазовой диаграмме. Исследование выполнено с использованием суперкомпьютерного комплекса НИУ ВШЭ [15].

Список литературы

1. M. C. Revelle, J. A. Fry, B. A. Olsen and R. G. Hulet, 1D to 3D Crossover of a Spin-Imbalanced Fermi Gas, *Phys. Rev. Lett.* 117 235301 (2016)
2. Y. Liao, A. S. C. Rittner, T. Paprotta et al., Spin-Imbalance in a One-Dimensional Fermi Gas, *Nature*, 467, 567–569 (2010)
3. R. Senaratne, D. Cavazos-Cavazos, S. Wang, F. He, Ya-Ting Chang, A. Kafle, Han Pu, Xi-Wen Guan, R. G. Hulet, Spin-charge separation in a 1D Fermi gas with tunable interactions, <https://arxiv.org/abs/2111.11545>
4. Fulde, P. & Ferrell, R. A. Superconductivity in a Strong Spin-Exchange Field. *Phys. Rev.* 135, A550–A563 (1964).
5. Larkin, A. I. and Ovchinnikov, Y. N. Inhomogeneous state of superconductors. *Sov. Phys. JETP* 20, 762-769 (1965).
6. K. Yang, Inhomogeneous superconducting state in quasi-one-dimensional systems, *Phys. Rev. B* 63, 140511(R) (2001)
7. G. Orso Attractive Fermi Gases with Unequal Spin Populations in Highly Elongated Traps *Phys. Rev. Lett.* 98, 070402 (2007)
8. B. Sundar, J. A. Fry, M. C. Revelle, R. G. Hulet and K.R. A. Hazzard, Spin-imbalanced ultracold Fermi

gases in a two-dimensional array of tubes, *Phys. Rev.* 102 033311 (2020)

9. A. Datta, K. Yang, A. Ghosal, Fulde-Ferrell-Larkin-Ovchinnikov state in strongly correlated d-wave superconductors, *Phys. Rev. B* 100, 035114 (2019)

10. U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Annals of Physics*, Volume 326, Issue 1, 2011, 96-192

11. E. Burovski., R. Ikhsanov, A. Kuznetsov, M.Yu. Kagan. Phase diagrams of polarized ultra-cold gases on attractive-U Hubbard ladders, International Conference on Computer Simulation and Beyond ed. L. Schur, *Journal of Physics Conference Series* 1163(5), UNSP 012046 (2019)

12. A.Potapova, I.Pile, E. Burovski., Quasi-one-Dimensional Polarized Superfluids: a DMRG Study, *Supercomputing 8th Russian Supercomputing Days, RuSCDays 2022, Moscow, Russia, September 26–27, 2022, Revised Selected Papers* Vol. 13708. Springer, 2022, P.303-312

13. G. Orso, E. Burovski, T.Jolicoeur Luttinger liquid of trimers in Fermi gases with unequal masses *Phys. Rev. Lett.* 104, 065301 (2010)

14. M. Fishman, S. R. White, E. M. Stoudenmire The ITensor Software Library for Tensor Network Calculations <https://arxiv.org/abs/2007.14822> (2020)

15. Kostenetskiy P.S., Chulkevich R.A., Kozyrev V.I. HPC Resources of the Higher School of Economics // *Journal of Physics: Conference Series*. 2021. Vol. 1740, No. 1. P. 012050. DOI: <https://doi.org/10.1088/1742-6596/1740/1/012050>.

ИСПОЛЬЗОВАНИЕ МОДЕЛИ СО СКОРОСТЬЮ ЗВУКА $\sqrt{2/3}$ В МЕТОДЕ РЕШЕТОЧНОГО УРАВНЕНИЯ БОЛЬЦМАНА НА ПРИМЕРЕ ДВИЖЕНИЯ ЧАСТИЦЫ В ПОТОКЕ ПУАЗЕЙЛЯ

Гуськова М.С.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В работе моделируется движение частицы в двумерном потоке Пуазейля методом решеточного уравнения Больцмана. Используются две модели: одна стандартная D2Q9 со скоростью звука $c_s = \sqrt{1/3}$, а вторая с другим набором скоростей и скоростью звука $c_s = \sqrt{2/3}$. Использование модели со скоростью звука $c_s = \sqrt{2/3}$ позволяет уменьшить время релаксации при прочих равных параметрах, что повышает точность. Отличия в результатах незначительны, если время релаксации τ близко единице.

Введение

Задачи микрофлюидики характеризуются маленьким значением числа Рейнольдса $Re = \frac{uL}{\nu}$, часто за счет большого значения кинематической вязкости. В решеточном методе Больцмана (LBM) кинематическая вязкость связана с временем релаксации τ , как $\nu = c_s^2(\tau - \frac{1}{2})$. Время релаксации входит в оператор столкновения и определяет, как быстро жидкость сходится к равновесному состоянию. При моделировании акустических задач τ должно быть как можно ближе к 0.5, для обычных задач τ выбирается близким к 1. Скорость звука в квадрате c_s^2 зависит от модели LBM, которая состоит из набора векторов скоростей и соответствующих им весов. Если увеличить скорость звука в квадрате в два раза, то вязкость тоже увеличится в два раза при том же времени релаксации. Движение частицы в потоке является известной задачей и часто становится тестовым для новых методов расчета взаимодействия частицы и потока [1,2]. Целью этой работы является исследование влияния скорости звука в модели LBM на движение частицы в течении Пуазейля.

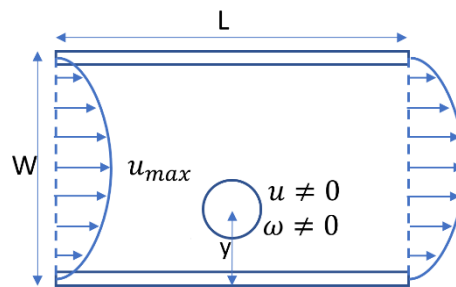


Рис. 1. Схема расчетной области

Течение Пуазейля – аналитическое решение уравнений Навье-Стокса в случае стационарного течения между двумя неподвижными плоскостями при наличии градиента давления [3]. Скорость в таком течении меняется вдоль толщины слоя жидкости по параболическому закону, достигая наибольшей величины в середине слоя.

Жесткая сферическая частица с нейтральной плавуучестью, двигаясь в течении Пуазейля, может сходиться к центру цилиндра или к некоторой оси между стенками и центральной осью. Этот эффект был описан в статье [4] и называется эффектом Сегре-Зильберберга. На частицу действуют: подъемная сила Саффмана, торможение из-за стенок, отраженные волны. В случае не центральной оси частица будет вращаться. Частица вблизи стенок будет двигаться медленнее потока. На ее линейную скорость, угловую скорость и итоговое положение влияет вязкость жидкости, число Рейнольдса, отношение диаметра частицы к диаметру канала, градиент скорости потока.

В статье [5] для двумерной задачи и простого сдвигового течения было определено положение равновесия для частицы в 0.5 расстояния от стенок до центральной оси и 0.6 для потока Пуазейля. Однако полученное приближение дает большие по-

грешности около стенок [6]. Частица нейтральной плавучести в течении Пуазейля сходится к $\beta^* = 0.19$ ($\beta^* = 0.81$), где $\beta = y/W$, где y расстояние между центром частицы и стенкой и W расстояние между стенками. Если частице запрещено вращаться, то $\beta^* = 0.26$ ($\beta^* = 0.74$). Для частицы с плавучестью, отличной от нейтральной, положение равновесия зависит от разницы плотностей частицы и жидкости и от того, может ли частица вращаться.

В работе [2] проведено численное исследование поведения частицы в потоке Пуазейля, получено эмпирическое выражение для подъемной силы. Положение равновесия сдвигается к стенке при росте числа Рейнольдса для каждого фиксированного \bar{a} ; и положение равновесия сдвигается к центральной оси при возрастании $\bar{a} = a/R$ для фиксированного числа Рейнольдса $\bar{U}_m = \frac{Re}{8\bar{a}^2}$, где a радиус частицы, R радиус канала.

В [1] предложен метод для расчета взаимодействия твердой частицы и жидкости, в качестве тестовой задачи исследуется эффект Сегре-Зильберберга для трех случаев: 1. двумерная задача, 2. трехмерная задача с прямоугольным каналом и 3. трехмерная задача с круглым каналом. Для фиксированного размера частицы положение равновесия h_e сдвигается к стенке при росте числа Рейнольдса и при фиксированном числе Рейнольдса положение равновесия сходится к центральной оси при возрастании размера частицы. Для двумерного случая показана зависимость h_e от y/W для $Re = 50$, где d/W , т.е. $y \in [0.2, 0.7]$.

Методы

Для моделирования потока используется метод решеточного уравнения Больцмана (Lattice Boltzmann method LBM). В LBM набор скоростей (векторов), каждому из которых соответствует вещественное число – популяция f_i [7]. На каждой итерации по времени происходит столкновение популяций и их перераспределение по векторам и этап распространения, когда популяции переходят в другой узел в соответствии со своим направлением.

Обычной модели D2Q9 соответствует набор скоростей \vec{c}_i : 1 скорость длины 0, 4 скорости длины 1 и 4 скорости длины $\sqrt{2}$. На Рис. 2 показаны векторы различной длины (R). Таким образом модели D2Q9 кроме нулевой скорости соответствуют векторы, заканчивающиеся на окружности голубого цвета и оранжевого. Плотность в каждой точке \vec{x} рассчитывается, как $\rho(\vec{x}) = \sum_i f_i$, а скорость, как сумма произведений $\rho\vec{u}(\vec{x}) = \sum_i f_i \cdot \vec{c}_i$. В работе используется оператор столкновений BGK (Бхатнагара-Гросса-Крука) $f_i = -1/\tau(f_i - f_i^{eq})$, где равновесные значения популяций определяются формулой $f_i^{eq} = \omega_i \rho (1 + \frac{\vec{u} \cdot \vec{c}_i}{c_s^2} + \frac{(\vec{u} \cdot \vec{c}_i)^2}{c_s^4} - \frac{\vec{u} \cdot \vec{u}}{2c_s^2})$.

Значения весов и наборы скоростей для двумерной модели со скоростью звука $c_s^2 = 2/3$ были получены с помощью алгоритма, реализованного в [8]. Этот набор скоростей показан в таблице 1. В отличие

от обычной модели не используются векторы длиной 1 (соответствующее значение $\omega_i = 0$), а используются векторы длиной 2.

Таблица 1. Векторы и их веса для моделей

Длина	Векторы	$c_s^2 = \frac{1}{3}$	$c_s^2 = \frac{2}{3}$
0	{0, 0}	4/9	4/9
1	{1, 0}, {0, 1}, {-1, 0}, {0, -1}	1/9	0
$\sqrt{2}$	{1, 1}, {-1, 1}, {-1, -1}, {1, -1}	1/36	1/9
2	{2, 0}, {0, 2}, {-2, 0}, {0, -2}	0	1/36

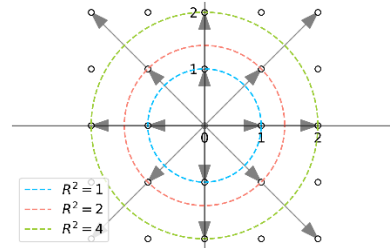


Рис. 2. Группы векторов разных длин для узла с координатами (0, 0)

Граничные условия сверху и снизу стенки bounce-back, скорость жидкости около стенок равна 0. Вдоль оси X периодические граничные условия. Для получения потока Пуазейля ко всем узлам расчетной решетки применяется внешняя сила $F = 8u_{\max}v/W^2$, $f_i^{in} = f_i^{in} + \omega_i \vec{c}_i \cdot \vec{F}/c_s^2$ [9].

Методом погруженной границы моделировалась частица [10]. Она представляется набором из точек на границе $N = 8$. Количество ibm-итераций 5. Частица движется в согласии с уравнениями движения классической механики, которые интегрируются явным методом Эйлера.

Для моделирования была написана программа на языке Python 3.8, код доступен в репозитории [11].

Результаты

Частица диаметром $D = 20$ движется в потоке Пуазейля с максимальной скоростью $u_{\max} = 0.03$ в двумерном канале длиной $L = 300$ и шириной 80 (Рис.1). Начало координат находится в левом нижнем углу расчетной области, ось X направлена вдоль направления потока, ось Y ей ортогональна. Сверху и снизу граничные условия без проскальзывания (стенки), вдоль оси X периодические граничные условия. Время релаксации τ , кинематическая вязкость ν и скорость звука c_s^2 варьируются.

На Рис.3 показана y-координата центра частицы в зависимости от времени. Для каждого набора параметров проводилось моделирование движения частицы из двух разных начальных положений: выше равновесного значения y_{st} $y_0 = \frac{y_c}{W} \approx 0.75$ и ниже $y_0 = \frac{y_c}{W} \approx 0.5$. Для различных начальных координат частица сходится к одному и тому же y_{st} (показаны полупрозрачной горизонтальной линией). Сплошные

линии голубая и оранжевая отличаются временем релаксации и вязкостью. Оранжевые линии отличаются временем релаксации, чем время τ меньше, тем больше y_{st} . Голубая линия и оранжевая пунктирная имеют разные значения вязкости, чем больше вязкость, тем ближе частица к центру канала.

На Рис.4 показаны результаты моделирования для параметров, где время релаксации близко к единице. Для модели с большей скоростью звука при той же вязкости можно получить меньшее время релаксации. Чем меньше время релаксации, тем точнее результат. Для τ близком к 1 различия в y_{st} уже не такие большие, как в случае с $\tau = 4.1$ и $\tau = 2.3$. Чем меньше вязкость, тем быстрее частица сходится к своему равновесному состоянию.

Заключение

На примере эффекта Сегре-Зильберберга было показано влияние времени релаксации и скорости звука на движение одиночной частицы в потоке Пуазейля. Использовались две двумерные модели LBM со скоростями звука $c_s = \sqrt{1/3}$ и $c_s = \sqrt{2/3}$. Оба набора скоростей позволили рассчитать равновесное положение частицы в потоке Пуазейля. При одинаковых вязкостях и временах релаксации близким к единице результаты близки. При одинаковом времени релаксации в модели с большей скоростью звука вязкость выше, а значит, число Рейнольдса меньше. Уменьшение числа Рейнольдса всего в два раза, а не на порядки, для многих задач не существенно, особенно по сравнению с затратами на реализацию новой модели. Однако модели с другой скоростью звука можно использовать для изучения влияния отраженных волн.

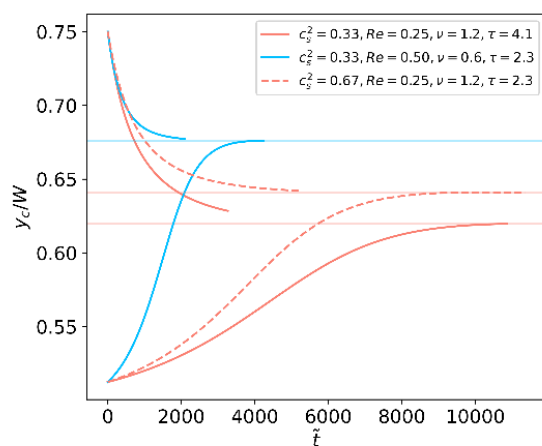


Рис. 3. у-координата центра частицы в зависимости от времени

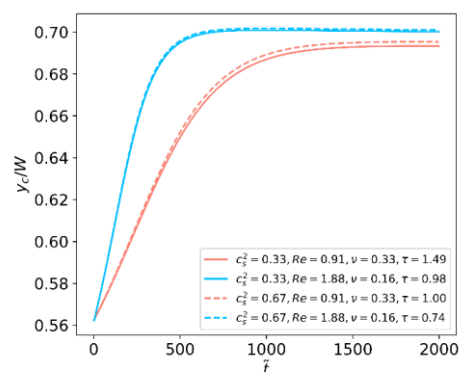


Рис. 4. у-координата центра частицы в зависимости от времени для τ близких к единице

Список литературы

1. Esipov D. V et al. Direct numerical simulation of the Segre-Silberberg effect using immersed boundary method // J. Fluids Eng. American Society of Mechanical Engineers Digital Collection, 2020. Vol. 142, № 11.
2. Yang B.H. et al. Migration of a sphere in tube flow // J. Fluid Mech. Cambridge University Press, 2005. Vol. 540. P. 109–131.
3. Landau L.D., Lifshitz E.M. Fluid Mechanics (Course of Theoretical Physics, v. VI). Pergamon Press, 1960.
4. Segre G., Silberberg A. Radial particle displacements in Poiseuille flow of suspensions // Nature. Springer, 1961. Vol. 189, № 4760. P. 209–210.
5. Ho B.P., Leal L. Inertial migration of rigid spheres in two-dimensional unidirectional flows // J. Fluid Mech. Cambridge University Press, 1974. Vol. 65, № 2. P. 365–400.
6. Vasseur P., Cox R.G. The lateral migration of a spherical particle in two-dimensional shear flows // J. Fluid Mech. Cambridge University Press, 1976. Vol. 78, № 2. P. 385–413.
7. Krüger T. et al. The lattice Boltzmann method // Springer Int. Publ. Springer, 2017. Vol. 10, № 978–3. P. 4–15.
8. Spiller D., Dünweg B. Semiautomatic construction of lattice Boltzmann models // Phys. Rev. E. APS, 2020. Vol. 101, № 4. P. 43310.
9. He X. et al. Analytic solutions of simple flows and analysis of nonslip boundary conditions for the lattice Boltzmann BGK model // J. Stat. Phys. Springer, 1997. Vol. 87, № 1. P. 115–136.
10. Inamuro T. Lattice Boltzmann methods for moving boundary flows // Fluid Dyn. Res. 2012. Vol. 44, № 2.
11. Гуськова М. GitHub-репозиторий lbm_cs2_2_3 [Electronic resource]. 2023. URL: https://github.com/msguskova/lbm_cs2_2_3.

Секция

«Информационно-коммуникационные технологии»

ПОИСК ИЗОБРАЖЕНИЙ НА МОБИЛЬНОМ ТЕЛЕФОНЕ (ОС ANDROID) С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Мотявин М.В., Чихватова А.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматривается процесс разработки программного решения, которое будет являться инструментом под операционную систему Android, способным осуществлять поиск изображений среди данных телефона по текстовому запросу пользователя, который отражает содержание изображений или объектов на них. Для решения данной задачи программное решение будет использовать методы машинного обучения, которые хорошо зарекомендовали себя для классификации и поиска по данным.

Введение

В условиях непрерывного роста данных необходима реализация решений способных осуществлять упрощение и автоматизацию повседневных процессов. Например, поиск по большому количеству медиа данных хранимых пользователями мобильных устройств. Для обработки и классификации данного типа данных, а также решения поставленных задач обычно используются методы машинного обучения.

Разрабатываемое программное решение с использованием машинного обучения планирует решать проблему долгого и затрудненного поиска изображений, которые хочет найти пользователь, среди изображений на мобильном устройстве. В ходе работы планируется реализация данного инструмента под операционную систему Android, как наиболее популярной и распространенной операционной системы среди мобильных устройств.

Для разработки программного решения необходимо провести: анализ литературы по тематике исследования, анализ аналогов разрабатываемого решения, анализ методов и способов разработки подобных решений.

Актуальность

В современном мире непрерывно растет поток данных производимых и потребляемых пользователями сети Интернет. По статистическим данным в 2018 году глобальный мобильный трафик данных составил 19,01 эксабайта в месяц. Ожидается, что к 2022 году мобильный трафик данных достигнет 77,5 эксабайт в месяц во всем мире при совокупном ежегодном темпе роста в 46 %. На мобильные устройства приходится примерно половина веб-трафика во всем мире. В первом квартале 2019 года мобильные

устройства (за исключением планшетов) генерировали 48,71 % глобального трафика веб-сайтов и сервисов [1].

Количество пользователей мобильных устройств продолжает расти в геометрической прогрессии. Более двух третей (67,1 %) населения мира в настоящее время пользуются мобильными телефонами, а число уникальных пользователей достигло 5,31 миллиарда к началу 2022 года. За 2021 год общее количество пользователей мобильной связи в мире выросло на 1,8 %, и с этого времени за 2021 год появилось 95 миллионов новых пользователей. Кроме того, аналитики «We Are Social» подсчитали, что средний пользователь использует мобильный телефон примерно пять часов в день. Это означает, что в 2022 году 5,3 млрд абонентов совокупно проведут более 1 млрд лет, используя мобильные телефоны [2]. Причина такого роста заключается в предоставлении обширных возможностей при использовании смартфона. Как и в случае с другими гаджетами возможности, предоставляемые мобильным устройством в первую очередь, облегчают выполнения повседневных задач, путем использования различных приложений, позволяющих решать огромный спектр пользовательских задач. Это, в свою очередь, порождает бурный рост количества исследований в области адаптации современных технологий к использованию на мобильных устройствах, включая как облачные сервисы, так и самостоятельные решения, работающие непосредственно на мобильном устройстве.

Наибольшее количество данных, потребляемых и производимых пользователями мобильного телефона, являются медиа данные, в частности фото данные. В настоящее время каждый смартфон имеет встроенное приложение «Галерея», в котором хранятся все изображения, полученные разным способом (сфотографированные на камеру, скачанные из социальных сетей, полученные по почте и т. д.). В 2023 году по статистическим данным у среднего пользователя на смартфоне хранится около 2000 фотографий [3]. Из-за такого объема изображений возникает проблема ориентирования и осуществления поиска среди фото-данных телефона. В телефонах отсутствует встроенная функция поиска нужного изображения по текстовому запросу, поэтому пользователю приходится вспоминать примерное время (год, месяц), когда было сделано или получено изображение и путем поиска (пролистывания) по всей галерее искать нужное фото, что является не самым удобным способом, который затрачивает, порой, большое количество времени на данную процедуру. Именно поэтому создание приложения, которое будет закрывать данную потребность, является актуальной темой для разработки.

Обоснование выбора ОС Android для разработки решения

Мобильное приложение – специальная программа, работающая на смартфонах и других мобильных устройствах с целью выполнения определенной задачи [4]. Большинство мобильных приложений ориентированы под операционную систему Android. Связано это с тем, что Android на в настоящее время является самой распространенной мобильной операционной системой.

По данным «Statcounter Global Stats» на ноябрь 2022 года доля Android на глобальном рынке мобильных операционных систем составляет 71,96 %, в то время как доля iOS составляет 27,48 %. Остальные мобильные операционные системы, такие как Samsung и KaiOS, вместе составляют менее 1 % доли рынка [5].

Говоря о рынке магазинов приложений в третьем квартале 2022 года, пользователи Android могли выбирать между 3,55 миллионами приложений в Google Play, что сделало Google Play магазином приложений с наибольшим количеством доступных приложений. Apple App Store был вторым по величине магазином приложений с примерно 1,6 миллионами доступных приложений для iOS [6].

Загрузить приложение для Android можно с любого веб-сайта, а не только из магазина Google Play. Помимо магазина Google Play, приложения могут размещаться на сторонних веб-сайтах. Комплект для разработки программного обеспечения Android позволяет это, а также возможность установить приложение на устройство напрямую или из командной строки.

Больше возможностей разработки приложений, взаимодействующих с аппаратным обеспечением устройства. В случае Android пользовательское ПЗУ (постоянное запоминающее устройство) относится к прошивке телефона. Поскольку Android имеет открытый исходный код, разработчик может редактировать и перекомпилировать код, а также адаптировать его для различных устройств.

Все это делает приложение доступным для большего количества пользователей и является основной причиной для инвестиций в платформу разработки приложений для Android.

Существующие инструменты по направлению исследования и их недостатки

В настоящее время существует несколько решений по тематике исследования. Одним из них является веб-сервис «Google Фото». Google Фото был представлен на конференции разработчиков Google I/O 2015, в тот же день была запущена веб-версия, а также приложения для iOS и Android. В этом сервисе реализована синхронизация снимков на всех устройствах. Можно осуществлять поиск по местам съемки, по некоторым текстовым запросам, а также разметка фотографий по людям, объектам, моментам и т. д.

Но у описанного выше инструмента есть ряд недостатков:

- чтобы установить данное приложение требуется поддержка смартфоном сервисов Google. Это

является проблемой ввиду зависимости от геополитики компании Google. Например, с 2019 года из-за санкций Google сервисы не работают на устройствах компании Huawei;

- для использования этого приложения пользователь должен иметь учетную запись в Google;

- ухудшение качества изображений. Чтобы занимать меньше пространства на серверах Google загруженные фото сжимаются, и при этом теряется оригинальное качество;

- необходимость постоянного наличия связи для выхода в Интернет для синхронизации с облаком является главным требованием;

- ограниченное бесплатное место на Google Диск. В каждом аккаунте Google по умолчанию доступно 15 ГБ для хранения данных. Но стоит отметить, что данное место используется также другими сервисами Google, такими как: «Google Диск», «Google Фото», «Gmail», а также другие сервисы, прикрепленные к аккаунту;

- нестабильная работа поиска. Изображения плохо классифицируются и поиск по некоторым базовым классам таким, как животные, например, собаки, не выдает результата, хотя в альбоме есть изображения данного класса;

- ограниченная конфиденциальность данного решения ввиду того, что в галерее у многих пользователей могут храниться фотографии документов, банковских карт или другие личные фотографии, но «Google Фото» будет производить синхронизацию и загрузку в облако всех изображений включая и конфиденциальные. При этом может возникать нежелательная проблема утечки данных.

Еще одним решением является функция «Поиск», которая появилась в галерее смартфонов с обновлением телефонов на новую графическую оболочку на базе Android 13. Например, данная функция доступна в графической оболочке OxygenOS 13 от компании OnePlus или MIUI 13 от Xiaomi. Это решение появилось и у некоторых других производителей.

Данная функция может производить поиск по некоторым ключевым словам, размечать лица и производить по ним поиск, производить поиск по месяцам и дате, а также местам, где было сделано фото. Недостатки решений в графических оболочках на базе Android 13:

- доступно только на устройствах, поддерживающих Android 13. Если устройство пользователя не соответствует стандартам и требованиям данной версии прошивки, то пользователь не сможет обновиться на новую графическую оболочку от компании-производителя смартфона и использовать данную функцию;

- весьма ограниченное количество классов, на которые могут быть размечены изображения, что уменьшает вариативность запросов и точность результата;

- поиск можно осуществлять только по одному ключевому слову, которое будет соответствовать названию встроенного класса, а поиск по словосочетанию или фразе не будет распознан;

– отсутствие обработки текстового запроса на наличие изображений, принадлежащих классу с синонимичным названием, а также обработки ошибок и опечаток в запросе.

Исходя из представленного анализа существующих решений, новизна реализуемого решения будет заключаться в том, что оно будет закрывать пробелы рассмотренных инструментов. Разрабатываемое решение будет иметь возможность поиска изображений по произвольным запросам пользователя, которые можно формулировать в виде предложений. Разрабатываемое решение не требует наличия доступа к сети Интернет, что делает решение более конфиденциальным, а возможность его использования независимым от внешней сети.

Планируемая реализация решения

Разрабатываемое мобильное приложение предназначено для поиска наиболее вероятных вариантов изображений среди изображений телефона по текстовому запросу пользователя, который описывает содержание изображения или его объектов. Предполагается, что пользователь при открытии приложения будет разрешать доступ к файлам телефона, для просмотра фотографий.

Для реализации данного решения будет использоваться предобученная модель CLIP от OpenAI [7]. Модель позволяет одновременно получать эмбединги по произвольному тексту и изображению – то есть с ней можно сравнивать текст и изображения в едином пространстве. Фактически это позволяет комбинировать категории и классифицировать изображения по более сложным текстовым описаниям. Модель является релевантной для решения поставленной задачи, так как совмещает в себе умение обучаться на тексте и изображениях одновременно, при этом не ограничивая себя определенными классами.

Принцип работы данной сети: в нейросеть подается изображение и текст, а она возвращает векторы изображения и текста. Затем можно посчитать косинусное расстояние и понять, насколько данный текст похож на изображение. В задаче классификации изображений можно выбирать класс по наибольшей близости векторов фотографии и текстового описания класса.

Для внедрения данной модели в мобильное приложение Android и последующего использования необходимо обученную модель конвертировать с помощью фреймворка Pytorch Mobile, который оптимизирует модели для мобильных устройств. После чего импортировать с помощью библиотек в модуль приложения и настроить развертывание на устройстве. В ходе конвертации получится модель с высокой производительностью, благодаря оптимизации модели и возможности использования встроенного аппаратного ускорителя в мобильное устройство.

Структурная схема работы приложения показана на рис. 1. При открытии приложения пользователь предоставляет доступ к галерее телефона. После этого извлекаются изображения, а также введенный пользователем текстовый запрос, которые кодируются с помощью описанной выше модели машинного

обучения, а также сопоставляются по семантическому сходству. Изображения с наибольшими оценками сходства возвращаются пользователю.

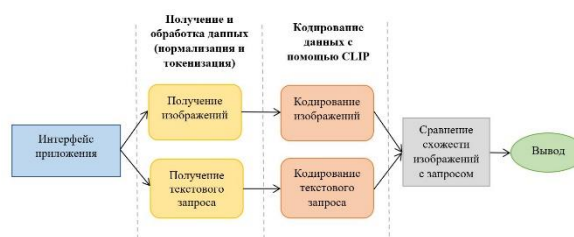


Рис. 1. Структурная схема мобильного приложения

Заключение

Тема данной работы является актуальной из-за большого постоянно растущего объема данных хранимых и производимых мобильными устройствами, которые требуют организации и навигации по ним. Планируемое решение будет решать одну из проблем ориентации по данным на мобильном устройстве – поиск изображений на смартфоне по текстовому запросу, делая процесс поиска изображений более быстрым, гибким, конфиденциальным и доступным вне зависимости от доступа к сети Интернет. Точность результатов запросов по тестовым данным составляет более 63 %. Для решения данной задачи будут использоваться методы машинного обучения, как одни из наиболее продвинутые и точные методов в области классификации и поиска данных.

Список литературы

1. Statista : [сайт]. – 2022. – URL: <https://www.statista.com/statistics/271405/global-mobile-data-traffic-forecast/> (дата обращения: 25.01.2023). – Текст : электронный.
2. Wearesocial : [сайт]. – 2022. – URL: <https://wearesocial.com/uk/blog/2022/01/digital-2022-another-year-of-bumper-growth-2/> (дата обращения: 25.01.2023). – Текст : электронный.
3. Photutorial : [сайт]. – 2023. – URL: <https://photutorial.com/photos-statistics/> (дата обращения: 25.01.2023). – Текст : электронный.
4. Анализ мобильных приложений и мобильного рынка / М. Г. Чубенко, Ю. А. Смирнова. – Текст : непосредственный // Вестник науки. – 2020. – Т. 1. – №. 6. – С. 32-36.
5. Statcounter Global Stats : [сайт]. – 2022. – URL: <https://gs.statcounter.com/os-market-share/mobile/worldwide> (дата обращения: 25.01.2023). – Текст : электронный.
6. Statista : [сайт]. – 2022. – URL: <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/> (дата обращения: 25.01.2023). – Текст : электронный.
7. OpenAI : [сайт]. – 2021. – URL: <https://openai.com/blog/clip/> (дата обращения: 25.01.2023). – Текст : электронный.

РАЗРАБОТКА ПРОГРАММЫ ГЕНЕРАЦИИ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ КОМБИНАЦИИ CNN, GAN И ДИФфуЗИОННЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Константинов А.С., Сидоров И.В.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматривается применения методов машинного обучения для генерации новых изображений. Разработанные решения применяются в задаче создания генератора логотипов в формате веб-приложения.

Введение

Актуальность темы работы обусловлена важностью исследования применения нейросетевых технологий для обработки изображений. В области генерации изображений по текстовому описанию 2022 год стал прорывным: появились коммерческие решения [1], а также решения с открытым исходным кодом [2]. Стремительное развитие технологий машинного обучения и нейронных сетей каждый день открывает новые возможности их использования. Так, например, глубинное обучение находит применение в дизайне [3] для генерации логотипов компаний. В настоящее время решения для создания логотипов зачастую ограничены использованием готовых форм, что приводит к однообразным результатам и не может послужить вспомогательным инструментом для дизайнеров и художников.

Предполагается, что дообучение диффузионной модели на работу с логотипами и ее объединение с другими архитектурами позволит достичь более высокого качества генерации изображений, а реализация итогового решения в виде приложения обеспечит удобство представления и использования итоговой системы.

В данной работе рассматриваются основные особенности и компоненты предлагаемого решения, а именно применение нейронных сетей, и архитектура веб-приложения, позволяющая получить отказоустойчивую и масштабируемую систему.

Существующие решения

Нейросетевые технологии давно применяются в дизайне, в частности, в сети существует множество решений, расширяющих инструментарий дизайнера логотипов. Например, генераторы логотипов, которые могут стать вдохновляющим фактором для художника в работе над новым проектом, предоставляя множество оригинальных вариантов.

Большинство существующих решений [4, 5] имеют схожий набор входных параметров, получаемых от пользователя: название организации, желаемая цветовая палитра, тематика работы организации (например, технологии или образование), стилистика

логотипа (например, плоский дизайн, использование натуральных форм). Подобный путь описания желаемого результата позволяет получить от пользователя достаточное количество информации для создания нового дизайна логотипа, однако практически все подобные сервисы имеют серьезный недостаток, который сказывается на оригинальности результатов – использование готовых форм и шаблонов от логотипа к логотипу, из-за чего идеи постоянно повторяются и трудно получить уникальный логотип.

Предполагается, что решить эту проблему можно путем использования диффузионных сетей.

Использование диффузионной модели

Задача генерации изображения по текстовому описанию с большим успехом получает новые решения в виде нейронных сетей в последние годы.

В рамках разрабатываемой системы предлагается преобразование пользовательского ввода, представляющего собой короткий опрос о желаемом результате (специализация организации, цветовые предпочтения и пр.), в текстовое описание, подаваемое на вход диффузионной нейронной сети Stable Diffusion, что позволяет получить структурированный и предсказуемый запрос для модели и, как следствие, более релевантный результат (см. пример на рис. 1).



Рис. 1. Пример генерации

Кроме того, можно улучшить качество выходных изображений, добавив вспомогательные элементы для модели. Во-первых, к ограниченному пользовательскому запросу предлагается добавлять заранее подобранные по результатам экспериментов с моделью особые слова, улучшающие генерацию. Во-вторых, пользовательский ввод будет обогащаться синонимами – для этого используется алгоритм Word2Vec, преобразующий слова в вектора, что позволяет математически оценить сходство между словами, что также улучшит качество генерации, позво-

для Stable Diffusion учесть наиболее важные аспекты желаемого результата. Наконец, дообучение модели на конкретную задачу, а именно создание логотипов, даст возможность получать более подходящие изображения по необходимой тематике.

Перенос стиля

Возможность переноса стиля с одного изображения на другое с сохранением исходного содержания впервые была представлена в 2015 году [6]. Преимущество данного подхода заключается в его простоте – используется предобученная на большом датасете ImageNet сверточная сеть и специальная функция потерь, состоящая из двух частей: одна оценивает разность в контенте между изображениями, а другая – разницу в стиле.

Иным подходом к переносу стиля с одного изображения на другое является использование генеративно-состязательных сетей (GAN) [7, 8]. Однако обучение и применение генеративных моделей является крайне ресурсозатратным и не дает заметных преимуществ в рассматриваемой задаче, поэтому для облегчения итогового решения было принято использовать первоначальный подход.

Для улучшения результата генерации изображения диффузионной моделью предлагается применить технологию замены стиля, давая возможность пользователю задать желаемый стиль результата.

Задача колоризации

Одной из дополнительных возможностей решения является выбор желаемой цветовой палитры результата. В задаче колоризации каждому пикселю целевого изображения присваивается определенный цвет. Методы колоризации можно грубо разделить на две категории: автоматическая и на основе пользовательских подсказок.

Первый тип не подходит под требуемый формат, так как требуется менять цвет на основе запроса, а не произвольно, поэтому необходимо использовать методы с применением подсказок. Колоризация с пользовательскими подсказками бывает нескольких видов: перенос цвета с референсного изображения, использование палитры цветов и использование линий или точек определенного цвета в нужных местах.

В предлагаемом решении используется модель iColoriT [9], которая производит окрашивание по подсказкам в виде точек. Модель является самым новым решением в данной области, имеет репозиторий с демонстрационным графическим интерфейсом, а окрашивание на этапе применения занимает доли секунды. В основе ее архитектуры лежит Vision Transformer, которые сейчас превосходят сверточные сети в области Computer Vision, поэтому iColoriT заметно превосходит предыдущие решения. Пример работы модели представлен на рис. 2.



Рис. 2. Пример колоризации

Генерация мокапов

Одним из преимуществ существующих решений и неотъемлемым элементом результатов работы над логотипом является создание мокапов (от англ. mock-up – макет) – примеров нанесения изображения на вывески, плакаты и различные объекты: футболки, кружки, блокноты и пр. Поэтому разрабатываемое решение предлагает такую возможность – сгенерированное изображение после всех настроек можно нанести на различные объекты для получения представления о внешнем виде логотипа в реальной среде его применения.

Техническая реализация

Программное решение представляет собой веб-сервис, создаваемый с использованием Python, JS и его производных и технологий контейнеризации и оркестрации для повышения отказоустойчивости. Одним из ключевых решений в программной реализации с точки зрения архитектуры (см. рис. 3) является применение очереди задач.

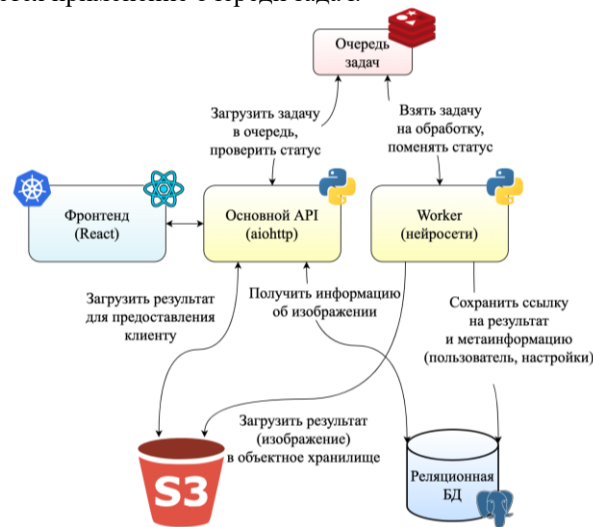


Рис. 3. Архитектура системы

Так как генерация и обработка изображений нейронными сетями являются ресурсоемкими блокирующими задачами, они не могут быть частью API веб-сервиса, чтобы не влиять на его доступность. Эти операции должны производиться асинхронно отдельными вычислительными машинами с доступом к графическому процессору – так называемыми worker'ами или обработчиками. Основной API системы только создает задачи на обработку в очереди,

реализованной на основе Redis. Обработчики берут на выполнение задачи из очереди и по завершении загружают результат в объектное хранилище, откуда главный API системы может их загрузить и предоставить клиенту.

Результаты

В результате работы создано отказоустойчивое и масштабируемое веб-приложение, позволяющее пользователям без специальных навыков создавать профессиональные макеты логотипов, обладающие намного большей оригинальностью и качеством, чем у аналогичных решений. Рассмотрим, например, создание логотипа пекарни с помощью сервиса Turbologo [5] (см. рис. 4) и с помощью предлагаемого в работе решения (см. рис. 5).

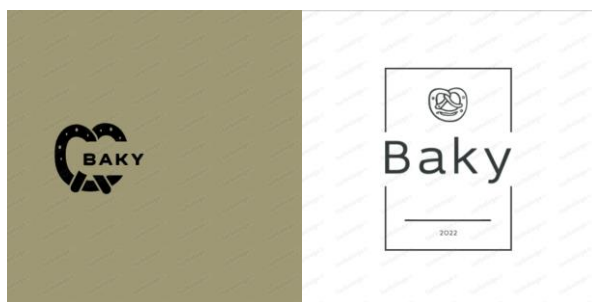


Рис. 4. Примеры логотипов, созданных с помощью Turbologo



Рис. 5. Примеры логотипов, созданных с помощью предлагаемого решения

Как можно видеть, существующие решения используют заготовленные формы, объединяемые различными способами в единый логотип, однако это происходит не всегда удачно. Кроме того, используемые формы зачастую повторяются.

Логотипы, получаемые с помощью разработанной системы, отличаются высокой оригинальностью

– генерируемые нейронной сетью объекты, а также их взаимное расположение, каждый раз уникальны.

Заключение

Внедрение искусственного интеллекта в различные сферы является актуальной задачей, поскольку новые технологии помогают упростить и ускорить существующие процессы, позволяя профессионалам сосредоточиться на более важных этапах работы. В ходе исследования были проанализированы и отобраны лучшие решения в сфере нейронных сетей, на основе которых было создано собственное приложение, позволяющее дизайнерам, компаниям и частным лицам создавать уникальные и высококачественные логотипы.

Предполагается, что предложенное решение может оказать влияние на индустрию дизайна, а также на научные открытия в области создания изображений. Исследователи могут использовать созданный в рамках работ набор данных и предобученную модель для дальнейшего совершенствования. Пользователи смогут экономить время на генерации идей и их воплощении с помощью веб-приложения.

Список литературы

1. DALL·E 2 - Open AI : [электронный ресурс]. - 2015. - URL: <https://openai.com/dall-e-2> (дата обращения: 20.01.2023).
2. Rombach R. et al. High-resolution image synthesis with latent diffusion models //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2022. – С. 10684-10695.
3. ART.LEBEDEV. Николай Иронов – дизайнер и нейросеть : [электронный ресурс]. - 2022. - URL: <https://ironov.artlebedev.com/ru> (дата обращения: 18.01.2023).
4. FreeLogoDesign : [электронный ресурс]. - 2023. - URL: <https://logo-maker.freelogodesign.org> (дата обращения: 21.01.2023).
5. Turbologo : [электронный ресурс]. - 2022. - URL: <https://turbologo.com> (дата обращения: 21.01.2023).
6. Gatys L. A., Ecker A. S., Bethge M. A neural algorithm of artistic style //arXiv preprint arXiv:1508.06576. – 2015.
7. Shahriar S. GAN Computers Generate Arts? A Survey on Visual Arts, Music, and Literary Text Generation using Generative Adversarial Network // Displays. 2021.
8. Mallika, Ubhi J.S., Aggarwal A.K. Neural Style Transfer for image within images and conditional GANs for destylization // J Vis Commun Image Represent. Academic Press Inc., 2022. № 85.
9. Yun J. et al. iColoriT: Towards Propagating Local Hints to the Right Region in Interactive Colorization by Leveraging Vision Transformer //Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. – 2023. – С. 1787-1796.

ИССЛЕДОВАНИЕ ОПТИМАЛЬНЫХ ГРАФОВ ЦИРКУЛЯНТНОЙ ТОПОЛОГИИ

Селин М.П.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматриваются параметры графов циркулянтной топологии применительно к их использованию в сетях на кристалле. Исследуются методы генерации графов и описывается собственное программное обеспечение. Производится сравнение параметров полученных графов с теоретически достижимыми характеристиками циркулянтов.

Введение

С развитием многопроцессорных систем, в том числе систем на кристалле и сетей на кристалле, для повышения производительности вычислительных систем необходимо осуществлять поиск способа соединения взаимодействующих элементов сети или системы, например, ядер. Существуют различные топологии сетей, которые имеют свои преимущества и недостатки. Одной из перспективных топологий являются циркулянты [1-2]. Циркулянты превосходят часто используемые топологии mesh [3] и torus [4] по диаметру, среднему расстоянию и ширине бисекции. Несмотря на свои параметры, циркулянты на данный момент не используются при проектировании реальных кристаллов СБИС.

Существует проблема генерации графов циркулянтной топологии. Для поиска оптимального графа необходимо выбрать лучший из всех существующих графов, что занимает значительное количество времени из-за большого количества возможных вариантов. Поэтому кроме оптимальных графов стоит рассматривать графы близкие к оптимальным, так как поиск такого графа может значительно сократить время вычислений при незначительном ухудшении параметров искомого графа.

Разработанное программное обеспечение

Нами создан генератор оптимальных циркулянтных топологий на языке C++. На вход приложения подается количество вершин и размерность графа, а также флаг поиска только среди кольцевых графов. На выходе генерируется .csv файл, содержащий количество вершин графа, набор образующих, диаметр, среднее расстояние, ширину бисекции и количество ребер.

Для решения задачи генерации циркулянтных графов применяется перебор различных вариантов циркулянтных графов. Для ускорения и оптимизации перебора используются следующие методы.

Программное обеспечение генерирует все наборы образующих, выполняющие неравенства (1) для поиска среди всех графов и (2) для поиска среди кольцевых графов:

$$1 \leq s_1 < s_2 < \dots < s_k \leq \left\lfloor \frac{n-1}{2} \right\rfloor \quad (1)$$

$$1 = s_1 < s_2 < \dots < s_k \leq \left\lfloor \frac{n-1}{2} \right\rfloor \quad (2)$$

Каждый набор образующих соответствует одному уникальному графу. Также проверяется условие (3):

$$\gcd(N, \{s_k\}) = 1. \quad (3)$$

Наборы, которые не соответствуют данному условию порождают графы, имеющие более 1 компоненты связности. Такие графы не могут рассматриваться как сеть из N узлов, поэтому дальнейшие вычисления для них не выполняются.

Для каждого набора выполняется алгоритм Дейкстры, возвращающий вектор расстояний до каждой вершины графа. Так как циркулянтные графы являются вершинно-транзитивными [5], алгоритм Дейкстры может выполняться из одной вершины, для оставшихся вершин результат будет идентичным.

В полученном векторе вычисляется диаметр графа и среднее расстояние в графе. Ширина бисекции вычисляется по формуле (4):

$$BW = \sum_{i=1}^k s_i. \quad (4)$$

Рассчитываемое по формуле (4) значение является верхней границей [6], так как существуют эквивалентные наборы образующих с различной шириной бисекции.

Графы ранжируются по диаметру, среднему расстоянию (по возрастанию) и ширине бисекции (по убыванию). В выходной файл попадают все графы с минимальным диаметром и средним расстоянием, ширина бисекции при этом не учитывается.

Сравнение реальных и теоретически достижимых параметров генерируемых графов

Теоретическое количество вершин, достигаемых в циркулянтном графе за i шагов вычисляется по формуле (5):

$$n_i \leq \sum_{j=1}^{\min(i,k)} 2^j \cdot C_k^j \cdot C_{m-1}^{j-1}. \quad (5)$$

Из данной формулы можно вывести наибольшее возможное количество вершин у графа с заданным диаметром d :

$$N_d \leq 1 + \sum_{i=1}^d n_i. \quad (6)$$

Из формулы (5) итеративно можно вычислить оптимальное среднее расстояние.

В качестве реальных графов используем графы размерностей 2 и 3, найденные разработанным программным обеспечением. В работе [7] доказано, что существует множество графов размерности 2, являющихся строго оптимальными, то есть достигающими теоретических значений диаметра и среднего расстояния. Выведена формула поиска образующих та-

ких графов. Среди графов размерности 2 будем рассматривать только кольцевые циркулянты. Для них это утверждение неверно, что видно из рисунка 1.

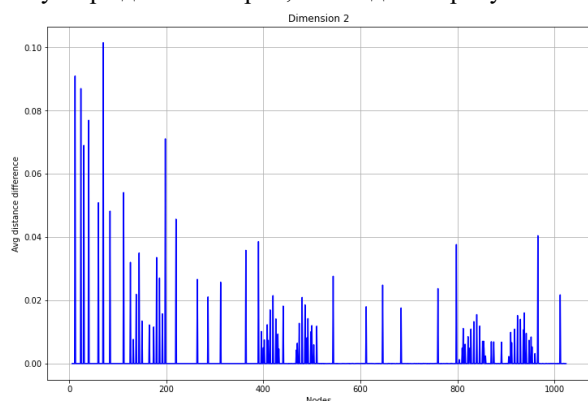


Рис. 1. Разность реального и теоретического среднего расстояния кольцевых графов размерности 2

С использованием тех же данных было вычислено, что диаметр на 1 больше теоретически достижимого имеют 44 кольцевых графа с количеством вершин от 6 до 1024. Среднее отклонение среднего расстояния в кольцевых графах от теоретического составило менее 0,002 хопа. Так как среди графов размерности 2 всегда существуют строго оптимальные вычисляемые за константное время, то поиск среди кольцевых графов не дает выигрыша во времени, но может ухудшить параметры итогового графа.

Для графов размерности 3 отклонения от теоретических показателей встречаются намного чаще. График разности реального и теоретически рассчитанного среднего расстояния приведен на рисунке 2.

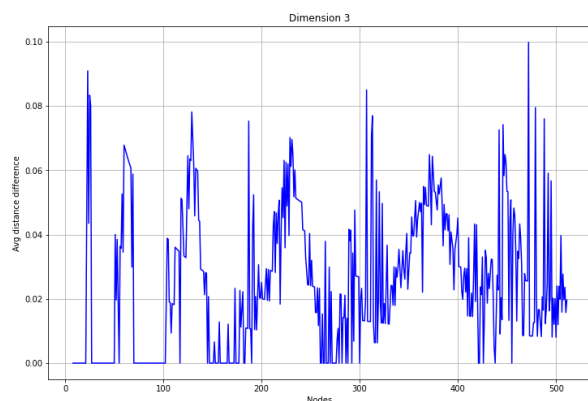


Рис. 2. Разность реального и теоретического среднего расстояния графов размерности 3

Среди графов размерности 3 не достигают теоретического диаметра 177 графов с количеством вершин от 8 до 511. Для кольцевых графов размерности 3 этот показатель составляет 187. Таким образом, всего для 10 топологий не существует кольцевого графа с диаметром, достижимым среди всех графов. Среднее отклонение среднего расстояния от теоретического составляет 0,0252 хопа для всех графов и 0,0249 для кольцевых. Данный результат приведен на рисунке 3.

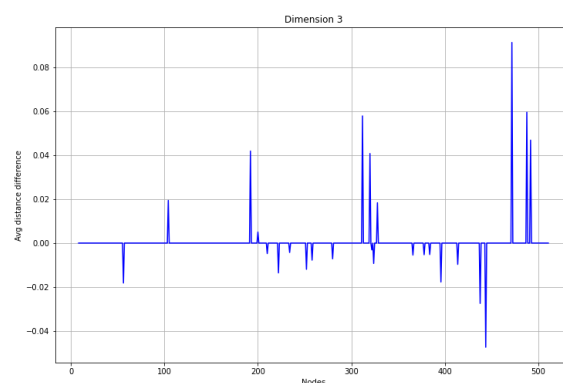


Рис. 3. Разность среднего расстояния кольцевых графов и всех графов размерности 3

Это оказывается возможным, так как среди 10 топологий с большим диаметром 9 имеют меньшее среднее расстояние. Кольцевой граф оставшейся топологии имеет больший диаметр и среднее расстояние, чем найденный среди всех. Для 15 топологий показатель среднего расстояния для кольцевых циркулянтов оказался хуже, чем среди всех графов при таком же значении диаметра.

Заключение

В работе были исследованы отклонения параметров циркулянтов, найденных среди кольцевых, от теоретически возможных оптимальных графов. Результаты показывают, что для циркулянтов размерности 2 поиск среди кольцевых графов не имеет смысла, так как приведет к ухудшению параметров и не приведет к улучшению времени поиска в случае поиска одного оптимального графа. Для циркулянтов размерности 3 наоборот, поиск среди кольцевых графов дает лучший результат по среднему расстоянию и лишь незначительно уступает по диаметру. Поиск кольцевых графов позволяет многократно сократить время вычислений. Например, для циркулянта на 512 вершинах с 4 образующими поиск кольцевого графа занимает 1 минуту, а поиск графа среди всех – немногим менее 2 часов. Такой выигрыш в скорости поиска может позволить использовать большие графы высоких размерностей при проектировании реальных кристаллов СБИС, так как их поиск не будет занимать недели.

Список литературы

1. Huang X., F Ramos A., Deng Y. Optimal circulant graphs as low-latency network topologies //The Journal of Supercomputing. – 2022. – P. 1-20.
2. Romanov A. Y., Amerikanov A. A., Lezhnev E. V. Analysis of approaches for synthesis of networks-on-chip by using circulant topologies //Journal of Physics: Conference Series. – IOP Publishing, 2018. – Vol. 1050. – No. 1. – P. 012071.
3. Rane U. V., Gad R. S., Panem C. Design of Network on Chip (NoC) computing node for mesh topology using soft-core NIOS-II processor //Journal of Physics: Conference Series. – IOP Publishing, 2021. – Vol. 1921. – No. 1. – P. 012075.
4. Punhani A., Faujdar N., Kumar S. Design and evaluation of cubic torus network-on-chip architecture

//International Journal of Innovative Technology and Exploring Engineering. – 2019. – Vol. 8. – No. 6. – P. 1672-1676.

5. Watkins M. E. Connectivity of transitive graphs //Journal of Combinatorial Theory. – 1970. – Vol. 8. – No. 1. – P. 23-29.

6. Монахова Э. А., Монахов О. Г. О некоторых характеристиках циркулянтных и тороидальных структур вычислительных систем //Вестник СибГУТИ. – 2013. – № 3 (23). – С. 63-69.

7. Beivide R. et al. Optimal distance networks of low degree for parallel computers //IEEE Transactions on Computers. – 1991. – Vol. 40. – No. 10. – P. 1109-1124.

АВТОМАТИЗАЦИЯ ВЕРИФИКАЦИИ СБИС С ПОМОЩЬЮ МЕТОДОВ ПОИСКА ОШИБКИ

Овчинникова А.А.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе предлагается автоматизированная система поиска причины ошибки программного теста на этапе верификации RTL-модели. Для этого описывается как сам подход, так и существующее окружение.

Введение

По мере того, как RTL-модели становятся все более и более нагруженными, сложность функциональной верификации возрастает в разы. Проблемы возникают в двух направлениях: увеличение затрачиваемого времени на выполнение проверки и неуверенность в ее полноте. Для решения вышеперечисленных проблем в последние годы в промышленности де-факто стала стандартом методология UVM, но даже при ее использовании важна также среда моделирования [1].

Верификация ASIC занимает от 60 до 70 процентов всего производственного цикла, и поэтому этой задаче в мире цифрового дизайна отводится специальное место. Чтобы сократить общее время выхода на рынок, крайне важно завершить этап верификации в наиболее короткое время [2].

В данной работе представлена автоматизированная система поиска причины ошибки программного теста. Эта система предоставляет специалисту дополнительные артефакты прохождения теста в удобном для восприятия формате, однако решение о причине ошибки остается за ним.

Система верификации дизайна

Дизайн, находящийся на этапе функциональной верификации, можно разделить на 3 основные составляющие: RTL-модель (на одном из HDL), функциональные тесты (на языке программирования C или SystemVerilog) и тестовое окружение (система,

обеспечивающая взаимодействие модели и тестов, формирующая результат прохождения тестов). Тестовое окружение является связующим элементом между дизайном и тестом. Как правило, оно состоит из множества верификационных блоков (Verification Intellectual Property, VIP) для используемых в дизайне протоколов, эталонных моделей для проверки выходных значений, блок проверки целостности данных(scoreboard), тестовых сценариев и программного кода для поддержания процесса симуляции тестов [3]. При этом источником ошибок может быть любой компонент (рис. 1).



Рис. 1. Источники программных ошибок на этапе верификации дизайна

Основным способом оценки полноты верификации дизайна является сбор покрытия. Метрики покрытия кода, переключения и функциональные метрики позволяют количественно оценить степень надежности компонента, вероятность его безошибочной работы [4]. Цель этапа верификации – получить наивысший возможный процент покрытия дизайна. При этом покрытие собирается только после успешного прохождения программного теста. В случае ошибки требуется найти и устранить причину.

Логическая диагностика комбинационных схем по своей сути является сложной задачей. Пространство решений растет экспоненциально с увеличением количества соединений цепи и количества неисправностей [5]. Это связано с тем, что реализация спецификации (HDL или неисправный чип) рассматривается как “черный ящик”, управляемый на первичных входах и наблюдаемый на первичных выходах [6]. Этот подход, являющийся традиционным, генерирует артефакты в виде трасс прохождения теста. Инженер по верификации проводит дальнейший анализ зачастую эмпирически и с высокой неопределенностью [4]. При этом для более полного анализа и систематизации данных часто требуются повторные запуски для получения дополнительных артефактов. Такой подход не только ненадежен, но и требует значительных усилий и времени. Для сокращения затрат ресурсов можно провести анализ с помощью автоматизированной системы.

Стоит заметить, что автоматизация поиска ошибки нетривиальная задача. Три вышеперечисленных составляющих дизайна, находящегося на этапе функциональной верификации, разрабатываются одновременно, соответственно, они нестабильны. При появлении ошибки не всегда можно точно понять, в каком из компонентов проблема.

Для упрощения задачи автоматизации поиска ошибки будем считать, что тестовое окружение работает стабильно. Приведем несколько причин, почему такое допущение возможно. Во-первых, при отладке инфраструктуры не используется спецификация, а значит оно не входит в цикл разработки дизайна, описанный выше, и оно не является частью конечного продукта. Ошибки тестового окружения возникают вне зависимости от типа верифицируемого блока. Благодаря этому их легче обнаружить, ведь в большинстве у каждого дизайна есть своя специфика при сбоях. При появлении массовых падений с одинаковым текстом ошибки можно с уверенностью отнести такой сбой к инфраструктурному. Во-вторых, для полной диагностики проблемы необходимо прибегать к дополнительному анализу процессов в системе, например сбоев в работе вычислительных кластеров. В этой работе используется информация из участка системы, ограниченного тестовым окружением и относящегося к циклу разработки и верификации дизайна.

Таким образом, область неопределенности сократилась до RTL-модели и программных тестов.

Автоматизированная система поиска ошибки

К поиску причины возникновения ошибки теста можно подойти с двух сторон. С одной стороны, данные о процессе прохождения теста могут указать на тот элемент, с которого началась поломка. С другой стороны, ошибка появилась как следствие проблемы логики, описанной в программном коде, а значит может быть найдена путем анализа внесенных в код изменений.

Стоит заметить, что в каждом из подходов, описанных выше, используется сравнение с эталоном: эталоном трассы прохождения теста и эталоном программного кода соответственно. Для достижения этого требуется ввести еще одно допущение – у системы поиска ошибки всегда есть референс. Этого можно легко добиться, например, тем, что использовать систему только для заведомо верно проходящих тестов.

Данные о прохождении тестов собираются тестовым окружением в виде различных артефактов. Среди них можно выделить:

- трассы процессов в тестовом окружении;
- трассы обращений по шинам;
- трассы процессов в ЦПУ.

Для начала анализа обязательно обратиться к трассе прохождения теста, где будут указаны сообщения об ошибке. Дополнить картину при необходимости можно используя остальные артефакты. Трассы обращений по шинам позволяют наглядно увидеть несоответствия изменений сигналов и могут локализовать ошибку вплоть до момента ее появления и неверно выставленного сигнала. Трассы процессов в ЦПУ предоставляют информацию об ошибках обращения к другим компонентам системы. Зная модуль, в котором возникла ошибка, можно связать ошибки в разных компонентах между собой.

Изменения, внесенные в код между версией, на которой тест проходил, и той, на которой он вывел

ошибку, представляют собой список потенциальных причин поломки. Такие изменения удобно отслеживать с помощью системы контроля версий. Этот список можно сузить, проассоциировав верифицируемые модули с перечнем описывающих их файлов. Дополнительным способом извлечения информации из этих данных является понятная для человека и машины маркировка изменений (коммитов), такое как семантическое версионирование и семантические коммиты.

Разработанная по приведенному описанию система сокращает рутинный анализ для инженера и передает эту работу машине. Она призвана извлечь гипотезы о причине ошибки из доступных артефактов, локализуя причину возникновения ошибки, сокращая тем самым время поиска ошибок. Инженер по верификации продолжает быть ответственным за окончательное решение и внесение правок, согласно выявленной неточности.

Заключение

Результат данной работы – это автоматизированная система поиска причины ошибки теста. По сути, работа состоит в добавлении дополнительных инструментов для анализа процессов в течение прохождения теста. Для этого используются артефакты тестового окружения и анализ изменений исходного кода.

Тестирование предложенного подхода еще не было проведено на практике. Результаты, ожидаемые после внедрения данного подхода:

- Экономия времени специалиста при поиске причины ошибки на 9%;
- Сокращение времени простоя задачи и воспроизведения ошибки на 8%;
- Уменьшение числа исполнителей, привлекаемых для выяснения причины падения теста на 10%.

Список литературы

1. Yang H. Y. Highly automated and efficient simulation environment with UVM //Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test. – IEEE, 2014. – p. 1-3.
2. Jain P., Shah M. V., Patel B. Automated Verification System for Functional Coverage Extraction //2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA). – IEEE, 2018. – p. 1870-1875.
3. Kim N. et al. How to automate millions lines of top-level UVM testbench and handle huge register classes //2012 International SoC Design Conference (ISOC). – IEEE, 2012. – p. 405-407.
4. Poulos Z., Veneris A. Clustering-based failure triage for rtl regression debugging //2014 International Test Conference. – IEEE, 2014. – p. 1-10.
5. Veneris A., Hajj I. N. Design error diagnosis and correction via test vector simulation //IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. – 1999. – Vol. 18. – No. 12. – p. 1803-1816.
6. Smith A. et al. Fault diagnosis and logic debugging using Boolean satisfiability //IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. – 2005. – Vol. 24. – No. 10. – p. 1606-1621.

ИССЛЕДОВАНИЕ ИЕРАРХИЧЕСКИХ ТОПОЛОГИЙ ДЛЯ ПРОЕКТИРОВАНИЯ СЕТЕЙ НА КРИСТАЛЛЕ

Рзаев Э.Р., Рыжов А.Р.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Данная работа является логическим продолжением предыдущих работ, связанных с исследованием перспективных топологий сетей на кристалле, основывающихся на прямом и корневом произведении циркулянтных графов. В этой статье мы описываем и сравниваем статические характеристики иерархического метода построений более сложных структур по сравнению с корневым и прямым произведениями графов. Также в данной работе представлен результат проведения анализа интеграции атомных структур (наименьших структурных единиц масштабируемых топологий) в различные уровни иерархической структуры. Представлено сравнение полученных структур с ранее исследованными структурами, являющимися результатами применения корневого произведения графов Пэли.

Введение

Основным направлением данной работы является изучение статических характеристик результата замещения топологий различных уровней иерархической структуры [1]–[3] циркулянтными графами вида $C(N; s_1, s_2, \dots, s_n)$ и их подклассами (графами Пэли [4]).

Данные структуры имеют высокую актуальность для проектирования топологий сетей на кристалле (СтнК), поскольку с возрастанием сложности вычислений, которая наблюдается в последние годы, следует повышение требований по производительности к подсистеме связи между вычислительными узлами. Однопроцессорные системы не могут в полной мере удовлетворить новым требованиям, поэтому необходим переход на более совершенные многопроцессорные системы. СтнК являются тем решением, к которому приходят разработчики при проектировании многопроцессорных систем. СтнК – это множество вычислительных модулей, объединенных общей подсистемой связи, состоящей из маршрутизаторов и соединений между ними [5],[6]. В настоящее время СтнК достигают таких размеров [7], [8], что их практически невозможно реализовать в одном тактовом домене [9], поэтому они зачастую разделены на отдельные кластеры, расположенные в разных тактовых доменах, для чего хорошо подходят иерархические топологии.

Таким образом, необходим поиск и синтез новых структур иерархических топологий СтнК, которые бы могли организовать их подсистему связи таким образом, чтобы она удовлетворяла новым требованиям. Следует также отметить, что такие топологии могут применяться не только в СтнК, но и, напри-

мер, в суперкомпьютерных кластерах, где тоже актуально использование иерархических топологий [9], [10].

В работе рассматриваются статические характеристики графов: средняя длина путей, диаметр, количество ребер, количество вершин, плотность, степени вершин.

Характеристики количества ребер и степени вершин занимают центральное место в задаче поиска подходящих топологий, так как данные параметры накладывают ограничения, обоснованные физическими свойствами структуры кристалла.

Обзор предметной области исследования

В работах [5],[6] подробно описана актуальность использования графов Пэли в качестве основы для проектирования СтнК. В ней рассмотрены и проанализированы различные модификации графов Пэли, в частности прямое [11] и корневое произведение графов. Но существует и другой подход к масштабированию структуры сети, известный как иерархические топологии [3]. Топология SCCN, рассматриваемая в приведенной статье, показала свою релевантность в сравнении с аналогами. Одним из ключевых свойств данного подхода является классификация уровней топологии и ее атомных узлов – структур, что являются неделимой единицей на данном уровне топологии.

Одной из важнейших проблем является поиск семейств оптимальных циркулянтных графов для количества узлов больше 100 и степени вершин больше 4 [12], что является основанием для проведения поиска масштабируемых топологий и исследования их характеристик.

Иерархическая топология с графом Пэли на 0 уровне

Инструментарий, используемый для сравнения обсуждаемых топологий, состоит из библиотек `python3 (networkx, matplotlib, numpy)`, а также программного обеспечения поиска оптимальных циркулянтных графов PCG Console [13].

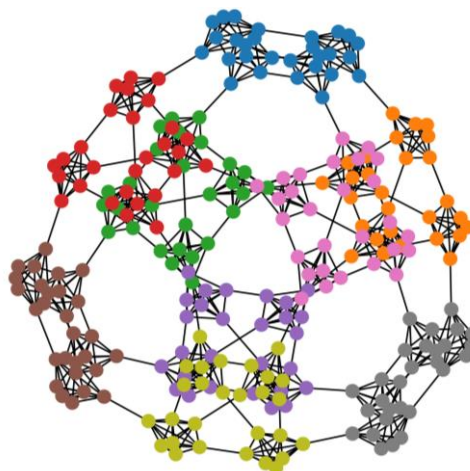


Рис. 1. Иерархическая топология с уровнями:
0 уровень – $QR(7)$, 1 уровень – соединение SCCN,
2 уровень – $torus(3,3)$

В статье [3] приводится описание построения иерархической топологии. В используемом методе построения иерархической топологии существует ограничение на количество топологий на нулевом уровне. Данный метод основывается на установлении связи между топологиями путем сдвига двоичных цифр индекса каждого шага узла влево или вправо. Данный метод сильно влияет на алгоритм маршрутизации, а следовательно, на динамические характеристики графов, которые не рассматриваются в данной статье.

На рисунке представлен $QR(7)$ на 0 уровне топологии, на 1 уровне представлено объединение 4 графов $QR(7)$ по описанному методу в приведенных выше статьях. Получившаяся структура помещается на следующий уровень в топологию $torus(3, 3)$.

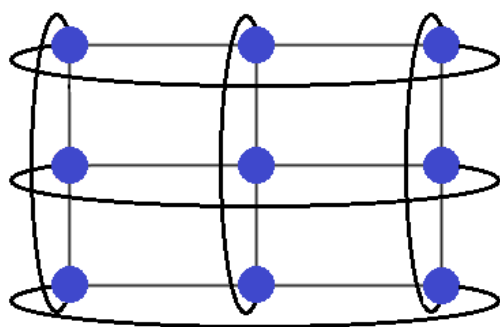


Рис. 2. Топология $torus(3,3)$

Иерархическая топология с графом Пэли на 2 уровне

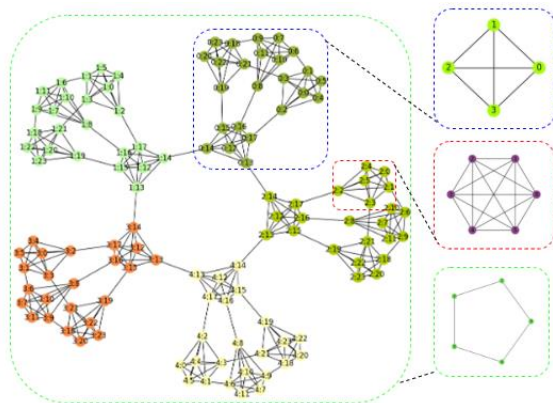


Рис. 3. Иерархическая структура с уровнями: 0 уровень – полносвязный граф на 6 вершинах, 1 уровень – соединение SCCN, 2 уровень – $QR(5)$

На рисунке 3 представлен альтернативный метод построения иерархической топологии. Данный метод также является результатом модификации первоначальной структуры SCCN путем интеграции графов Пэли, но на 2-й уровень.

На рисунке пунктирными линиями выделены структуры, справа к которым приводятся топологии данного уровня (отображение между вершинами используемой топологией каждого уровня и кластерами подграфов является биективным), объясняющая логику принципа соединения дочерних графов на каждом из приведенных уровней.

Анализ графов Пэли в иерархической структуре

В качестве сравнения в прошлой статье [5] изучался метод корневого произведения графов Пэли между собой до 101 вершины в результирующем графе. Результирующий граф – граф, являющийся результатом произведения двух графов G и H .

В случае иерархической структуры можно однозначно произвести биекцию структур по количеству вершин, использующихся в графах Пэли (до 97 вершин включительно в графе Пэли). В случае корневого произведения графов Пэли $G \times H$ объем получившейся выборки данных имеет большую размерность. Данный факт представляет трудность для объекта сравнения между собой обоих методов. Поэтому было решено аппроксимировать получившиеся параметры путем их объединения и усреднения по группам количества вершин графа G , то есть графа, стоящего на первом месте в произведении.

Ниже представлены статистические характеристики графов, полученных в результате применения методов, описанных выше пунктах, а именно: граф Пэли на 0 уровне – Paley0, граф Пэли на 2 уровне – Paley2, корневое произведение графов Пэли – Root_QRQR.

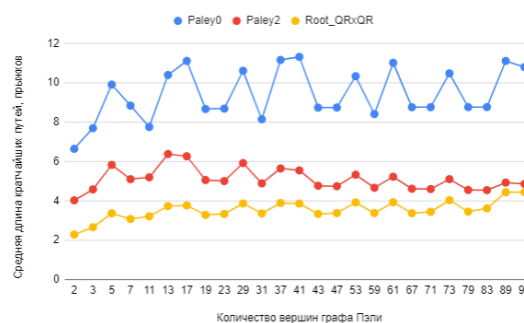


Рис. 4. Средняя длина путей

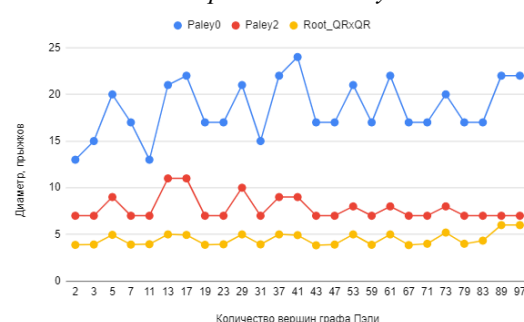


Рис. 5. Диаметр

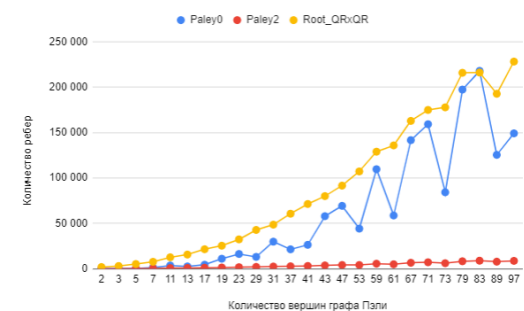


Рис. 6. Количество ребер

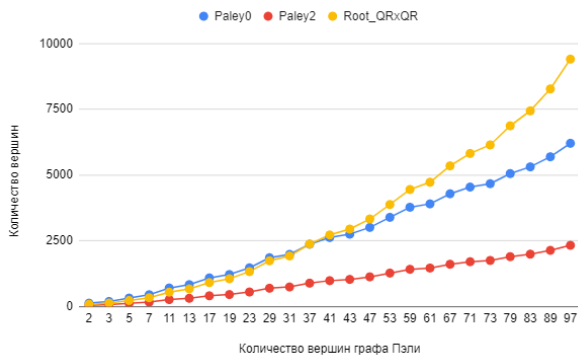


Рис. 7. Количество вершин

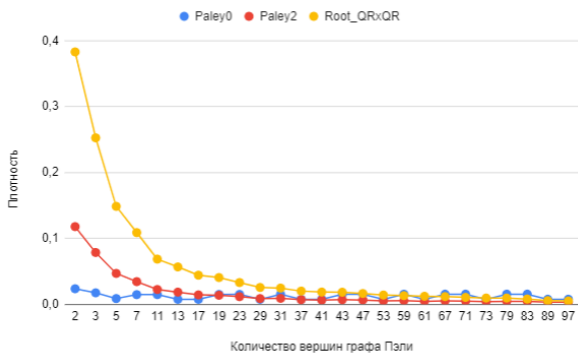


Рис. 8. Плотность

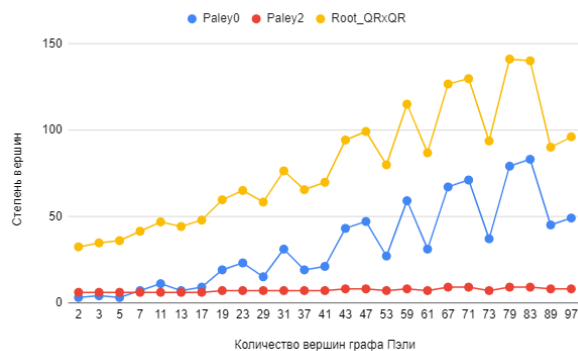


Рис. 9. Степени вершин

Таким образом, метод N копий структуры полного графа на 6 вершинах через граф Пэли (Paley2) оказывается лучшим по всем параметрам, по сравнению с замещением mesh. В случаях, где данный метод уступает корневому произведению графов Пэли, различие небольшое. Стоит отметить, что скорость увеличения количества ребер очень хорошо компенсирует описанный выше аспект.

Заключение

В данной статье были разобраны методы замещения различных уровней иерархической структуры с проведением анализа статических характеристик графов. В частности, были рассмотрены графы Пэли и результаты их интеграции в исследуемую топологию. Анализ замещения показал, что новая структура имеет конкурентоспособные параметры по сравнению с исходной структурой, построенной на полном графе из 6 вершин.

Список литературы

1. M. N. M. Ali, M. M. Hafizur Rahman, D. K. Behera, and Y. Inoguchi, "Static Cost-Effective Analysis of a Shifted Completely Connected Network," *Advances in Intelligent Systems and Computing*, vol. 711, pp. 165–175, 2019, doi: 10.1007/978-981-10-8055-5_16.
2. M. N. M. Ali, M. M. H. Rahman, and T. M. Tengku Sembok, "SCCN: a cost effective hierarchical interconnection network for network-on-chip," *International Journal of Advancements in Computing Technology (IJACT)*, vol. 8, no. 5, pp. 70–79, 2016.
3. M. N. M. Ali, M. M. H. Rahman, A. A. Ibrahim, Y. Inoguchi, and E. Hossain, "The Static Performance Effect of Hybrid- Hierarchical Interconnection by Shifted Completely Connected Network," *IEEE Access*, vol. 9, pp. 99249–99265, 2021, doi: 10.1109/ACCESS.2021.3095146.
4. T. K. Kim and C. E. Praeger, "On generalised Paley graphs and their automorphism groups," *Michigan Mathematical Journal*, vol. 58, no. 1, pp. 293–308, 2009.
5. E. R. Rzaev and A. Y. Romanov, "The New Promising Network-on-Chip Topologies Development Using Product Operation," *Proceedings - 2021 International Russian Automation Conference, RusAutoCon 2021*, pp. 421–425, Sep. 2021, doi: 10.1109/RUSAUTOCON52004.2021.9537317.
6. Рзаев Э. Р., Романов А. Ю. Исследование перспективных топологий сетей на кристалле: применение корневого и прямого произведений графов Пэли // В кн.: Проблемы разработки перспективных микро- и наноэлектронных систем – 2021 (МЭС-2021) / Под общ. ред.: А. Л. Стемпковский. Вып. 1. ИПИМ РАН, 2021. С. 9-16., doi: 10.31114/2078-7707-2021-1-9-16.
7. D. Ditzel *et al.*, "Accelerating ML recommendation with over a thousand RISC-V/Tensor processors on esperanto's ET-SoC-1 Chip," *2021 IEEE Hot Chips 33 Symposium, HCS 2021*, Aug. 2021, doi: 10.1109/HCS52781.2021.9566904.
8. K. Rocki *et al.*, "Fast stencil-code computation on a wafer-scale processor," *International Conference for High Performance Computing, Networking, Storage and Analysis, SC*, vol. 2020-November, Nov. 2020, doi: 10.1109/SC41405.2020.00062.
9. A. Sheibanyrad, A. Greiner, and I. Miro-Panades, "Multisynchronous and Fully Asynchronous NoCs for GALS Architectures," *IEEE Design and Test*, vol. 25, no. 06, pp. 572–580, Nov. 2008, doi: 10.1109/MDT.2008.167.
10. F. al Faisal, M. M. Hafizur Rahman, and Y. Inoguchi, "HFBN: An Energy Efficient High Performance Hierarchical Interconnection Network for Exascale Supercomputer," *IEEE Access*, vol. 10, pp. 3088–3104, 2022, doi: 10.1109/ACCESS.2021.3138828.
11. X. Huang, A. F. Ramos, and Y. Deng, "Optimal circulant graphs as low-latency network topologies," *arXiv preprint arXiv:2201.01342*, 2022.

12. O. G. Monakhov and E. A. Monakhova, "Regular Network Class Features Enhancement Using an Evolutionary Synthesis Algorithm," *Science & Education of Bauman MSTU/Nauka i Obrazovanie of Bauman MSTU*, no. 10, 2014.

13. A.Yu. Romanov, A.Yu. Glukhikh "circulant-Graphs/PCG Console.exe," [Online]. Available: https://github.com/RomeoMe5/circulantGraphs/blob/master/PCG_Console/PCG_Console.exe.

МЕТОДИКА СОЗДАНИЯ ДАТАСЕТА АВТОМОБИЛЬНЫХ АВАРИЙ

Анджушева М.М., Бичурина Е.П.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В данной работе представлен комплекс программных средств для сбора, обработки, разметки и формирования датасета, состоящего из видеофрагментов автомобильных аварий. Работа проводилась с целью упрощения и автоматизации процесса создания датасета видео аварий. В рамках разработки методики предпринята попытка наиболее полно исследовать существующие решения, определить тенденции, важные для практического применения датасетов, и выявить их возможные недостатки. Комплекс предлагает стратегию для загрузки, разметки и хранения видео в датасете для практического применения датасета дорожных происшествий в задачах машинного обучения.

Введение

Безопасность на дорогах – это важная проблема, которая касается, без преувеличения, каждого. Одним из прогрессивных решений данной проблемы является применение методов машинного обучения. Так, многие авторы в работах [1] - [15] принимают попытки извлечь важную информацию дорожных сценариев, представляя множество направлений применения этих данных.

В настоящей работе исследована структура датасетов дорожных инцидентов, их актуальность и обоснование выбранной формы представления, способов сбора, разметки и хранения данных и представлена усовершенствованная методика с комплексом действий по созданию видеодатасета дорожных аварий.

Цель разработанной методики - повысить эффективность использования датасетов, содержащих видеофрагменты автомобильных аварий.

Задачи исследования:

1. Общий обзор анализ актуальных публикаций по видеодатасетам дорожно-транспортных происшествий (ДТП).

2. Выделение набора релевантных решений.
 3. Описание способов сбора данных, подходов к разметке видео, формы представления записей датасета.
 4. Изучение логики принятия решений по процессу сбора датасета.
 5. Изучение инструментов, используемых для подготовки и разметки видеозаписей.
 6. Поиск решения по хранению датасета и взаимодействию с облачным хранилищем.
 7. Изучение видов и способов разметки видео аварий.
 8. Разработка системы классификации видео аварий.
 9. Анализ платформ для размещения датасета.
 10. Разработка алгоритма действий по созданию датасета с видео ДТП с приложенным комплексом программных средств.
- В данной работе описан поэтапный план создания и реализации видеодатасета с учетом как условий временной разметки, так и метаданных.

Существующие решения

В датасетах дорожных инцидентов выделяются 3 основные группы по способу работы с данными: датасеты в формате CSV, датасеты с текстовым описанием видео и визуальные датасеты. Визуальные датасеты дорожных сценариев, главным образом, решают следующие задачи:

1. Предсказание аварий [3, 4].
2. Управление беспилотным автомобилем. [10, 12, 15].
3. Установление причинно-следственных связей дорожных инцидентов [6, 9, 11, 13, 14].

Датасеты, представленные в CSV формате [1], представляют из себя табличные данные о дорожных происшествиях. В таких датасетах важен лишь факт самого инцидента и подробное его описание по заранее выбранным признакам. Другой популярный подход нацелен на поиск взаимосвязи между видео и текстовым описанием происходящего [2]. То же самое можно проделать с видеозаписями дорожных происшествий [11] или, как вариант, добавить отрывок из сводки новостей о конкретной аварии.

Наиболее продвинутым подходом считается извлечение информации непосредственно из видео [3] - [15]. В настоящее время существует большое количество нейросетевых моделей, позволяющих извлекать признаки из изображений и видео [10, 15], решать задачи сегментации, детекции и много другое.

Изучив большое количество статей [1] - [15], можно сделать однозначный вывод, что создание датасетов дорожных происшествий востребовано и актуально. В ходе изучения существующих датасетов извлечены важные особенности сбора, разметки и хранения данных.

Датасет YouTubeCrash [3] включает в себя записи с видеорегистраторов, каждая из которых поделена на две части: аварийная и неаварийная сцена. В датасете Street Accident [4] к видеозаписям с видеорегистраторов вручную добавлены данные об адресе или GPS локации. Датасет Car Accident Detection and

Prediction (CADP) [5] содержит видео, разметка которых состояла из этапов временной и сегментации и пространственно-временной разметки. Система, предложенная в датасете BDD: 100K [6], включает в себя указание сцены (шоссе, городская улица, двор и т.д.), детекцию боксами, сегментацию, указание времени суток, разметку дороги. Каждое видео содержит также информацию о GPS/IMU, чтобы фиксировать траекторию движения.

Видео для The Cityscapes Dataset [7], прорывного датасета в задаче семантической сегментации, снимались специально для датасета и дополнялись информацией о температуре воздухе и GPS метками. Самый известный и прорывной датасет в задаче семантической сегментации городских сцен CamVid Dataset [8] содержит видео с семантическими метками класса. Mapillary Vistas Dataset [9] содержит 25000 сегментированных изображений высокого качества, взятые с www.mapillary.com – ресурс для публикации фото улиц со всего мира. D2-city Dataset [10] содержит видео с видеорегистраторов в городах Китая, взятых с платформы DiDi (сервис такси и каршеринга). Не выбирались последовательные фрагменты: если был выбран текущий клип, то следующий за ним фрагмент исключался. DeepRide Dataset [11] – это первый набор текстовых описаний для записей с видеорегистраторов. Он соединяет область автономного вождения с описанием видео, генерируя текстовое описание визуальной информации, видимой с помощью видеорегистратора. В ROAD Dataset [12] данные основываются на маршрутах беспилотного автомобиля Nissan Leaf и включают в себя видео с 6 камер, данные GPS и другую специфическую информацию с сенсоров автомобиля. Для аннотации использовался VOTT Tool, позволяющий копировать настройки разметки (bounding boxes и их метки) и применять их к следующему фрагменту.

Датасет DADA-2000 [13] призван решить задачу driving attention. Кроме того, авторы убеждены, что датасет позволит предсказывать аварию, а в будущем и вовсе устанавливать причинно-следственные связи произошедшей аварии. Для сбора данных использовались следующие платформы: Youtube, Youku, Bilibili, iQiyi, Tencent. Кроме временной разметки датасет содержит пространственную локализацию участников сцены. На основе временной и пространственной аннотации строится attention map. Для датасета DR(eye)VE Dataset [14] видео получали из специального эксперимента, в котором 7 мужчин и 1 женщина ездили на машинах с двумя камерами и сенсорами в течении 2 месяцев. Использовалась система SMI ETG 2w Eye Tracking Glasses, позволяющая отслеживать движение зрачков, изменение положения головы и т.д. Кроме того, использовались камера на крыше и видео с очков водителя. Датасет Car Crash [15] содержит записи аварий с видеорегистраторов с YouTube, с помощью нейросети VGG-16 были извлечены признаки и получены баундинг боксы для каждого видео, каждое видео было размечено с указанием времени суток, погодных условий, уча-

ствия машины, с которой производится съемка, в аварии.

В виду новизны предметной области, единый стандарт по сбору и разметке датасетов отсутствует. Так, способ разбиения видео на фрагменты в каждой работе свой, он основан на эвристиках и не обосновывается авторами (колеблется от 10 fps до 50 fps). Размечаются видео так же по-разному: в основном, производится детекция и/или сегментация участвующих в сценах объектах, применяются разные подходы к временной разметке и выделению момента происшествия, добавляются разные метаданные (погодные условия, время суток, тип дороги).

В таблице 1 приведен сравнительный анализ платформ для размещения датасетов.

Таблица 1. Платформы для размещения датасетов

	Zenodo	Mendeley Data	IEEE Dataport	Kaggle	Paperswith code
Цена	бесплатно	бесплатно	1 950 \$	бесплатно	бесплатно
DOI	есть	есть	есть	есть	нет
Популярность	2.5 млн/мес	0.34 млн/мес	0.14 млн/мес	9 млн/мес	1.4 млн/мес
Количество датасетов	170 тыс	10 млн	13.5 тыс	180 тыс	7.3 тыс
Соревнования по ML	нет	есть	есть	есть	нет
Форматы загрузки данных	любой	любой	любой	любой	любой
Возможность обновления датасета	нет	есть	есть	есть	есть

Алгоритм создания датасета аварий

Алгоритм создания видеодатасета дорожных аварий от идеи до реализации представлен на рисунке 1.

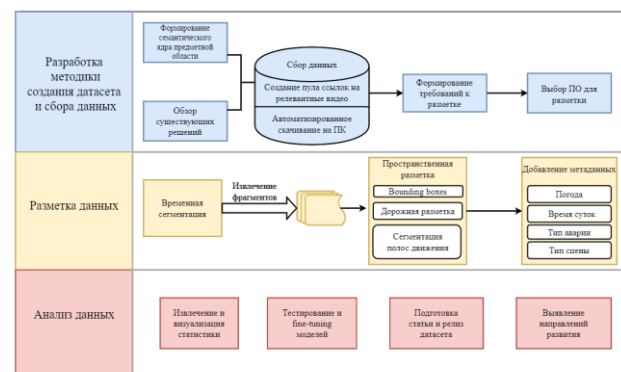


Рис. 1. Алгоритм создания датасета

Методика создания датасета и сбор данных осуществляется следующими действиями:

1. Убедиться в понимании предметной области и ознакомиться с уже имеющимися решениями.
2. Определить источники, откуда будут взяты релевантные данные. Скачать и подготовить видеоролики.
3. Сформировать требования к разметке данных.

4. Выбрать и установить программное обеспечение для разметки.

В изученных публикациях разметка как правило заключалась только в детекции или только в сегментации - авторы объясняют это трудоемкостью и дороговизной данной работы. Проанализировав предметную область, предлагается следующая разметка. Детекция транспортных средств, дорожных знаков, светофоров, пешеходов. Выделение дорожной разметки. Сегментация полос движения. Добавление метаданных (погода, время суток, тип дорожной сцены и др).

Разметка данных производится в несколько этапов:

1. Извлечение фрагментов посредством временной сегментации.
2. Пространственная разметка (bounding boxes, дорожная разметка, сегментация полос движения).
3. Добавление метаданных (погода, время суток, тип аварии, тип сцены).

Полученные результаты позволяют провести независимый анализ данных, содержащихся в датасете:

1. Извлечение и визуализация статистики.
2. Обучение и настройка ML-моделей на основе разработанного датасета.
3. Подготовка научной статьи и размещение датасета на специальной платформе.
4. Выявление направлений развития.

Экспериментальная проверка методики заключалась в создании датасета аварий на более 1000 записей. Осуществлен сбор материалов из открытых источников, произведено извлечение фрагментов с моментами аварий в установленном формате. Фрагменты размечены с выведением всех объектов в рекомендованном программном обеспечении, в аннотации к разметке добавлены метаданные. Подготовленные видео хранятся в упорядоченном виде в общем хранилище. Производится анализ датасета и подготовка к опубликованию собранного датасета на открытой платформе. Результаты эксперимента показывают, что методика является эффективной и удовлетворяет требованиям, предъявляемым к созданию датасета..

Заключение

Разработана методика создания видеодатасета ДТП, основанная на анализе существующих решений. Методология включает в себя этапы поиска видео, сбора, хранения и обработки. Подробно рассматриваются способы и варианты разметки видеофрагментов с использованием программного обеспечения. Предусмотрен анализ полученных данных, представлены варианты размещения результатов работы.

Список литературы

1. M. S., "US Accidents," 2019. [Online]. Available: https://smoosavi.org/datasets/us_accidents. [Accessed 1 november 2022]
2. Rovšek V., Batista M., Bogunović B. Identifying the key risk factors of traffic accident injury severity on

Slovenian roads using a non-parametric classification tree //Transport. – 2017. – Vol. 32. – No. 3. – P. 272-281.

3. Aditya A. et al. Collision Detection: An Improved Deep Learning Approach Using SENet and ResNext //2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC). – IEEE, 2021. – P. 2075-2082.

4. Chan F. et al. Anticipating Accidents in Dashcam Videos //2016 Computer Vision - ACCV 2016. – P. 136-153.

5. Shah A. P. et al. CADP: A novel dataset for CCTV traffic camera based accident analysis //2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). – IEEE, 2018. – P. 1-9.

6. Yu F. et al. Bdd100k: A diverse driving dataset for heterogeneous multitask learning //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2020. – P. 2636-2645.

7. Cordts M. et al. The cityscapes dataset for semantic urban scene understanding //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 3213-3223.

8. "CamVid Dataset," [Online]. Available: <http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/>.

9. Neuhold G. et al. The mapillary vistas dataset for semantic understanding of street scenes //Proceedings of the IEEE international conference on computer vision. – 2017. – P. 4990-4999.

10. Che Z. et al. D²-City: A Large-Scale Dashcam Video Dataset of Diverse Traffic Scenarios // arXiv preprint arXiv:1904.01975. – 2019. [Online]. Available: <https://arxiv.org/pdf/1904.01975.pdf>. [Accessed 1 november 2022].

11. Rafiq G. et al. DeepRide: Dashcam Video Description Dataset for Autonomous Vehicle Location-Aware Trip Description // IEEE Access. – 2022. – Vol. 10. – P. 107361-107375.

12. Singh G. et al. Road: The road event awareness dataset for autonomous driving //arXiv preprint arXiv:2102.11585. – 2021. [Online]. Available: <https://arxiv.org/pdf/2102.11585.pdf>. [Accessed 1 november 2022].

13. Fang J. et al. Dada-2000: Can driving accident be predicted by driver attention? analyzed by a benchmark // 2019 IEEE Intelligent Transportation Systems Conference (ITSC). – IEEE, 2019. – P. 4303-4309.

14. Palazzi A. et al. Predicting the Driver's Focus of Attention: the DR (eye) VE Project // IEEE transactions on pattern analysis and machine intelligence. – 2018. – Vol. 41. – No. 7. – P. 1720-1733.

15. "CarCrashDataset," [Online]. Available: <https://github.com/Cogito2012/CarCrashDataset#overview>. [Accessed 1 november 2022]

СИСТЕМА АВТОМАТИЗИРОВАННОГО РАСЧЕТА СТАЦИОНАРНЫХ ТЕПЛОВЫХ РЕЖИМОВ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

*Баканов Г.П., Подкопаева П.О.,
Цыгикало П.Н., Савинов А.И.*

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе предлагается решение задачи – расчета стационарных тепловых режимов интегральных микросхем (далее – ИМС) и трехмерная визуализация тепловой модели. Спроектирован программный комплекс с гибкой архитектурой расчётного ядра с возможностями внедрения, апробации и сравнения эффективности различных методов и алгоритмов в решении тепловых задач.

Введение

Существует ряд коммерческих инструментов теплового анализа, широко используемых в микроэлектронной промышленности: ANSYS, SolidWorks, COMSOL Multiphysics, и др. Однако, универсальность инструментов анализа подобных систем усложняет процесс создания геометрических моделей интегральных микросхем, содержащих большое количество элементов, требующих сопряжения, и задание граничных условий для расчета тепловой модели. Среди отечественного программного обеспечения, позволяющего осуществлять тепловое моделирование ИМС стоит отметить специализированный пакет АСОНИКА-Т, но отсутствие API затрудняет его использование для решения исследовательских задач. Все вышеперечисленные системы имеют проприетарные форматы тепловой модели и закрытый исходный программный код, что затрудняет проведение исследований, требующих оптимизации расчетных алгоритмов, или задание сложных пользовательских моделей материала. Целью работы является создание гибкого инструмента моделирования тепловых режимов ИМС с открытым исходным кодом, удобным форматом входной геометрии, а также возможностью настройки параметров визуализации тепловой модели. В работе учитывается опыт МИЭМ в разработке инструмента теплового анализа ИМС – квази-трехмерного симулятора Overheat[1] и является продолжением работы[2].

Следующие разделы освещают: постановку решаемой задачи, метод решения, описание этапов реализации метода, архитектуру системы и пример моделирования.

Постановка задачи

Задача теплового моделирования состоит в расчете распределения температуры в трехмерном твердом теле, занимающим объём V , ограниченную поверхностью S . Материалы считаются однородными и изотропными. Уравнение теплопроводности в декартовой системе координат имеет вид (1):

$$\lambda \frac{\partial^2 T}{\partial x^2} + \lambda \frac{\partial^2 T}{\partial y^2} + \lambda \frac{\partial^2 T}{\partial z^2} + Q = 0 \quad (1)$$

где $T = T(x, y, z)$ описывает распределение температуры (T – температура тела в точке (x, y, z)); λ – теплопроводность размерности $Вт/м \cdot ^\circ C$; $Q = Q(x, y, z, T)$ – интенсивность источников тепла внутри тела ($Вт/м^3$).

На поверхности $S = s_1 \cup s_2 \cup s_3$ заданы условия обмена с окружающей средой:

$$\lambda \bar{n} \frac{\partial T}{\partial n} \Big|_{s_1} = q; \quad \lambda \bar{n} \frac{\partial T}{\partial n} \Big|_{s_2} = \alpha(T - T_{cp}); \quad T|_{s_3} = T_{гр} \quad (2)$$

где α – коэффициент конвективного теплообмена ($Вт/м^2 \cdot ^\circ C$), q – интенсивность теплового потока ($Вт/м^2$), а T_{cp} – температура окружающей среды. На поверхности s_3 задается постоянная температура $T_{гр}$, конвективный теплообмен – на s_2 и тепловой поток – на s_1 .

Автоматизация решения задачи возможна путем машинной реализации численных методов. Метод конечных элементов (далее – МКЭ) является группой методов, в основе которых аппроксимация непрерывной функции на конечном числе подобластей (элементов). МКЭ является универсальным[3] для решения задач математической физики и его модификации применяются[4] в инструментах теплового анализа лидеров индустрии систем автоматизированного проектирования и разработки (далее – САПР) (ANSYS Discovery Live, Dassault Systèmes SolidWorks Flow Simulation).

Конечноэлементное решение задачи состоит из следующих основных этапов:

- Дискретизация области решения,
- Построение системы алгебраических уравнений на построенной КЭ сетке,
- Решение полученной системы уравнений,
- Визуализация результатов.

Дискретизация области решения

При реализации МКЭ применимы различные типы конечных элементов: тетраэдр, призма, параллелепипед и др. Наиболее часто встречающимися[5] типом является тетраэдр в силу преимуществ дальнейших вычислений, связанных с линейностью базисных функций и простой геометрии элемента. Существует обширное количество готовых библиотек с возможностями построения и геометрической обработки треугольных сеток: libigl, cgal[6], GeometryCentral или VCG. OpenMesh[7] и PMP[8] имеют несколько более широкую область применения и могут обрабатывать общие полигональные поверхности. Для трехмерных объектов, части libigl и GeometryCentral предлагают рудиментарную поддержку определенных твердотельных элементов, таких как тетраэдры или гексаэдры, но большая часть библиотеки сосредоточена на поверхностях. И наоборот, OpenVolumeMesh полностью ориентирован на объемы и может работать с общими многогранными элементами, но не поддерживает поверх-

ностные сетки. Только Geogram имеет унифицированную структуру данных, которая может содержать как элементы поверхности, так и элементы объемных фигур, но поддерживает только шестигранники, тетраэдры, призмы и пирамиды в качестве объемных ячеек. В отличие от всех этих альтернатив, Cinolib[9] имеет уникальную структуру данных, предназначенную для размещения любого типа поверхности и объемного элемента. В качестве бэкенда для построения тетраэдров алгоритмом Делоне используется TetGen[10] генератор. Таким образом для разбиения пространства была выбрана библиотека Cinolib.

Вычисление матрицы теплопроводности и вектора нагрузки

МКЭ позволяет свести исходную задачу о решении системы уравнений в частных производных (1) к задаче о решении системы линейных алгебраических уравнений (СЛАУ). Для построения этой системы функция температуры внутри каждого тетраэдрального конечного элемента с вершинами i, j, k, l аппроксимируется линейной функцией (3):

$$\begin{aligned} \Phi^{(e)} &= N_i^{(e)} \Phi_i + N_j^{(e)} \Phi_j + N_k^{(e)} \Phi_k + N_l^{(e)} \Phi_l = \\ &= [N_i^{(e)} \ N_j^{(e)} \ N_k^{(e)} \ N_l^{(e)}] \begin{Bmatrix} \Phi_i \\ \Phi_j \\ \Phi_k \\ \Phi_l \end{Bmatrix} = [N^{(e)}] \{\Phi\}, \end{aligned} \quad (3)$$

где $\Phi_i, \Phi_j, \Phi_k, \Phi_l$ – узловые значения функции температуры, N_i, N_j, N_k, N_l – базисные функции элемента, имеющие вид (4):

$$N_i^{(e)} = \frac{1}{6V^{(e)}} (a_i^{(e)} + b_i^{(e)}x + c_i^{(e)}y + d_i^{(e)}z), \quad (4)$$

где коэффициентами $a_i^{(e)} - d_i^{(e)}$ являются соответствующие строке узла i алгебраические дополнения матрицы координат вершин элемента (5):

$$\begin{bmatrix} 1 & x_i & y_i & z_i \\ 1 & x_j & y_j & z_j \\ 1 & x_k & y_k & z_k \\ 1 & x_l & y_l & z_l \end{bmatrix} \quad (5)$$

Вид базисных функций для узлов j, k, l аналогичен.

Известно, что решение уравнения (1) с граничными условиями (2) эквивалентно отысканию минимума функционала (6):

$$\begin{aligned} J &= \int_V \frac{1}{2} \left(\lambda \left[\left(\frac{\partial T}{\partial x} \right)^2 + \left(\frac{\partial T}{\partial y} \right)^2 + \left(\frac{\partial T}{\partial z} \right)^2 \right] - 2QT \right) dV + \\ &+ \int_{S_1} qT dS + \int_{S_2} \frac{\alpha}{2} (T - T_{cp})^2 dS \end{aligned} \quad (6)$$

Минимизация функционала (6) на множестве узловых значений функции T сводит задачу к следующей системе линейных уравнений:

$$[K]\{T\} + \{F\} = 0, \quad (7)$$

где $[K]$ – матрица теплопроводности:

$$[K] = \sum_{e=1}^E [K^{(e)}], \quad (8)$$

и $\{F\}$ – вектор нагрузки:

$$\{F\} = - \sum_{e=1}^E \{F^{(e)}\} \quad (9)$$

Матрицы (9) и (10) формируются суммированием вкладов отдельных элементов (10) и (13) соответственно:

$$\begin{aligned} [K^{(e)}] &= \int_{V^{(e)}} [B^{(e)}]^T [D^{(e)}] [B^{(e)}] dV + \\ &+ \int_{S_2^{(e)}} \alpha^{(e)} [N^{(e)}]^T [N^{(e)}] dS, \end{aligned} \quad (10)$$

где $[D^{(e)}]$ – локальная матрица теплопроводностей элемента (11), а $[B^{(e)}]$ – локальная матрица коэффициентов (12) элемента, полученных при дифференцировании базисных функций $[N^{(e)}]$ по x, y, z :

$$D^{(e)} = \begin{bmatrix} \lambda^{(e)} & 0 & 0 \\ 0 & \lambda^{(e)} & 0 \\ 0 & 0 & \lambda^{(e)} \end{bmatrix}, \quad (11)$$

$$[B^{(e)}] = \frac{1}{6V^{(e)}} \begin{bmatrix} b_i & b_j & b_k & b_l \\ c_i & c_j & c_k & c_l \\ d_i & d_j & d_k & d_l \end{bmatrix}, \quad (12)$$

$$\begin{aligned} \{F^{(e)}\} &= \int_{S_1^{(e)}} q^{(e)} [N^{(e)}]^T dS - \int_{V^{(e)}} Q^{(e)} [N^{(e)}]^T dV - \\ &- \int_{S_2^{(e)}} \alpha^{(e)} T_{cp} [N^{(e)}]^T dS, \end{aligned} \quad (13)$$

Для тетраэдрального элемента интегралы в выражениях (11) и (14) вычисляются аналитически в L-координатах. Результат вычисления и суммирования интегралов – числовые значения локальных матрицы теплопроводности $[K^{(e)}]$ и вектора нагрузки элемента $\{F^{(e)}\}$. С точки зрения программной реализации значения из локальных матрицы теплопроводности и вектора нагрузки записываются в матрицу $[K]$ и вектор $\{F\}$ глобальной системы уравнений (8) в соответствии с нумерацией узлов конечно-элементной сетки. Таким образом решение уравнения (1) с граничными условиями (2) сводится к системе линейных уравнений (7).

Решение системы линейных уравнений

Матрица теплопроводности системы (7) является ленточной и разреженной. Универсальным способом решения СЛАУ является метод последовательного

исключения (Гаусса), однако он не учитывает разреженность матриц.

Другой известный метод, который является модификацией метода Гаусса – LU-разложение. Исходная матрица представляется как произведение двух матриц: L (нижняя треугольная) и U (верхняя треугольная). Однако быстродействие данного метода неприемлемо низкое при решении больших СЛАУ (при 10000 неизвестных вычисление занимает порядка 3 часов), а также требуется дополнительное пространство для хранения[11]. Таким образом, данный метод является не очень эффективным.

Для хранения разреженных матриц наиболее популярными методами являются CSR (Compressed Sparse Row) и его модификация CSC (Compressed Sparse Column)[12].

В работе[13] предлагаются некоторые методы решения СЛАУ, специализированные под особенности метода конечных элементов. Данные методы оптимизируют структуру хранения разреженных ленточных матриц посредством подструктурирования или специального разбиения на основе взаимосвязи матрицы и ячеистой структуры. При этом предлагается приводить матрицу к блочной структуре, что может уменьшить количество операций и снизить затраты памяти.

Для решения СЛАУ применимы готовые библиотеки линейной алгебры (например, boost), однако они не учитывают специфические особенности матриц.

Планируется реализовать модификацию алгоритма прямого метода решения СЛАУ с помощью LU-разложения (Холецкого)[14]. Хранение матриц с помощью CSC позволяет упростить доступ к ненулевым элементам матрицы.

Архитектура программного комплекса

Языком разработки является C++ для эффективного выполнения операций с множеством элементов геометрии и перерасчёта параметров в реальном времени при взаимодействии с трёхмерной моделью. Обработка и визуализация геометрии реализуется на основе широко распространенного в индустрии САПР стандарта OpenGL, позволяющего достичь высокой производительности.

Архитектура программного комплекса поддерживает возможность интеграции дополнительных расчётных ядер. В дальнейшем данную возможность планируется использовать в исследовательской деятельности для сравнения, модификации и улучшения методов решения тепловых задач. В разрабатываемом программном комплексе входные данные геометрии, физических констант и граничные условия задаются в текстовом файле в соответствии с инструкцией к симулятору Overheat[1].

Предлагаемая архитектура программного комплекса представлена на рис. 1.

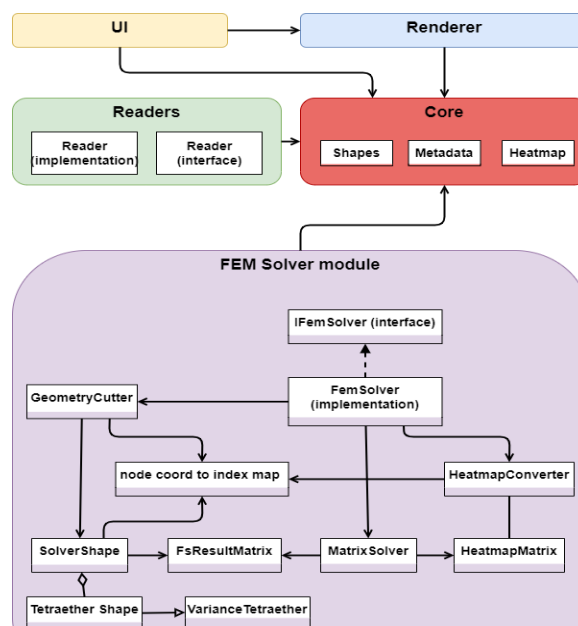


Рис. 1. Архитектурная схема программного комплекса

Программа принимает на вход файлы геометрии ИМС. Затем модуль Readers осуществляет считывание данных, и они передаются в хранилище в модуль ядра программы Core. Затем модуль расчёта температурного распределения Fem Solver принимает метаданные и геометрию из модуля Core и передаёт в подмодуль GeometryCutter для разбиения объема трехмерной геометрии на конечные элементы. GeometryCutter создаёт конечно-элементную сетку и возвращает конечные элементы SolverShape и структуру NodeCoordToIndexMap, хранящую координаты их вершин. Далее коэффициенты аппроксимирующих функций рассчитываются для каждого конечного элемента в TetraetherShape, а также происходит вычисление коэффициентов матрицы теплопроводности и вектора нагрузки. Матрицы глобальной системы линейных уравнений хранятся в структуре FsResultMatrix. MatrixSolver реализует алгоритм решения системы линейных уравнений с разреженной ленточной матрицей теплопроводности. Результатом работы MatrixSolver является массив температур в узлах Temperatures. Далее происходит интерполяция массива температур в набор тепловых карт по слоям (HeatmapConverter). Затем в модуле Renderer происходит преобразование температур, тепловых карт и их загрузка в сцену рендерера. В нем по полученной геометрии с помощью OpenGL создаются и повторно используются необходимые буферы вершин и текстуры для отображения карт температурного распределения, цвет которых на элементах корректно интерполируется при плотности сетки температурных значений меньше, чем плотность пиксельной сетки.

При разработке программного комплекса предполагается установить ряд ограничений на использование инструмента конечными пользователями и будущими разработчиками. Среди технических требований: процессор Intel core i7 11 поколения и выше, 16 гигабайт ОЗУ и более. Оперативная система Win-

dows 7 и новее или Linux. Также предполагается в ходе разработки документировать список всех библиотек и внешних модулей, которые интегрируются в систему в руководствах к разрабатываемому программному обеспечению. Для корректной работы программного комплекса требуется строгое следование требованиям и инструкциям из руководств пользователя и разработчика.

Пример моделирования

В качестве примера демонстрации ожидаемой работы программного комплекса был визуализирован (рис. 2) результат расчета стационарного теплового режима интегрального стабилизатора напряжения K142ЕН9 в симуляторе Overheat[1], разработанного группой сотрудников МИЭМ.

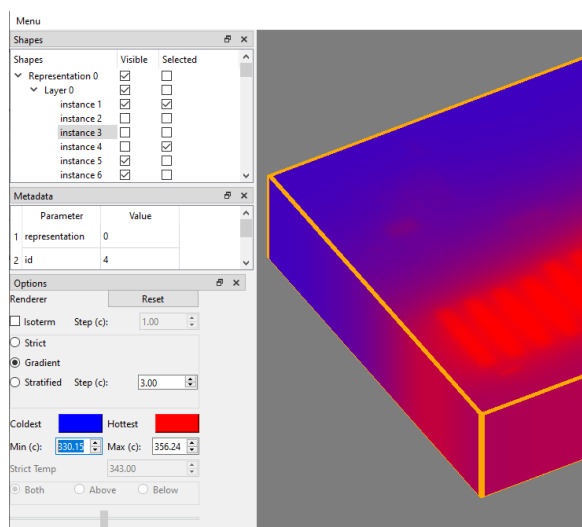


Рис. 2. Пример моделирования к 22.01.2023.

Программа позволяет взаимодействовать с трёхмерной моделью во вкладке Options: изменять координату точки обзора, угол обзора, настраивать цветовой диапазон, включать отображение изотерм с заданным шагом в °C, а также устанавливать градиентную или стратифицированную отрисовку с заданным шагом в °C. Иерархический список всех элементов предлагает управлять отображением и подсветкой граней любого элемента модели (окно Shapes).

Заключение

Спроектирован программный комплекс для решения задачи расчета тепловых режимов ИМС. Разрабатывается гибкое конечно-элементное ядро с возможностями модификации текущих и интеграции новых алгоритмов и методов решения тепловых задач. Разработан постпроцессор, визуализирующий тепловую модель и совместимый с ядром Overheat[1].

В дальнейшем планируется интеграция программного комплекса в библиотеку инструментов решения задач математической физики (пример инструмента библиотеки – [15]) факультета МИЭМ НИУ ВШЭ. Создание единой библиотеки инструментов позволит удовлетворить потребности научных групп, специализирующихся на исследованиях в

области решения задач математической и повысить технологическую независимость университета. Также появление открытого и гибкого инструмента имеет потенциал образовательной платформы для развития компетенций профильных специалистов в области разработки САПР.

Список литературы

1. Petrosjanc K.O., Ryabov N.I., Kharitonov I.A., Kozynko P.A. Electro-thermal simulation: a new Subsystem in Mentor Graphics IC Design Flow // EDA Publishing/THERMINIC, 2009. - с. 71–74.
2. Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов имени Е.В. Арменского. Материалы конференции. - М. ~: МИЭМ НИУ ВШЭ, 2022. - с. 218-220 - ISBN 978-5-903650-71-2. - Текст : непосредственный.
3. Huebner K., Dewhirst, D и др. The finite element method for engineers. //John Wiley & Sons, 2001 – с.3-14.
4. Magomedov I.A., Sebaeva Z.S. Comparative study of finite element analysis software packages // J Phys Conf Ser. IOP Publishing, 2020. - Vol. 1515 - с. 32073.
5. Серерлинд Л. Применение метода конечных элементов. //Мир, 1979. – с. 8–20.
6. Goren N., Fogel E., Halperin D. CGAL Made More Accessible.// arXiv:2202.13889v1 - 2022.
7. McCormack J., Porter B., Wetter J. A et al. Scriptable, generative modelling system for dynamic 3D meshes.// arXiv:2004.10354v1- 2012.
8. Wen X. et al. PMP-Net++: Point Cloud Completion by Transformer-Enhanced Multi-step Point Moving Paths.// IEEE transactions on pattern analysis and machine intelligence. - Vol. 45 - No. 1 – 2023.
9. Livesu M. Cinolib: A Generic Programming Header Only C++ Library for Processing Polygonal and Polyhedral Meshes // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, 2019. - Vol. 11820. - с. 64–76.
10. Si H. TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator // ACM Transactions on Mathematical Software (TOMS), 2015. - Vol. 41 – No.2.
11. Peter M. Development and Analysis of Sparse Matrix Concepts for Finite Element Approximation on general Cells, 2004. - Vol. 10.
12. Pissanetzky S. Sparse Matrix Technology. // ACADEMIC PRESS INC, 1984 – с.38-45.
13. George A. Sparse Matrix Aspects of the Finite Element Method. // Computing Methods in Applied Sciences and Engineering, 1976. - с. 3–22.
14. Davis T.A. Direct Methods for Sparse Linear Systems // Society for Industrial and Applied Mathematics, 2006. – с. 135-144.
15. Проект #831 [Электронный ресурс]. - URL: <https://cabinet.miem.hse.ru/#/project/831/> (дата обращения: 20.01.2023).

РАЗРАБОТКА МЕТОДИКИ И АЛГОРИТМОВ АВТОМАТИЗИРОВАННОГО ТЕСТИРОВАНИЯ ГРАФИЧЕСКИХ ПРИЛОЖЕНИЙ НА АППАРАТНО-ПРОГРАММНОЙ ПЛАТФОРМЕ «ЭЛЬБРУС»

Киреенко Л.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ,
АО «МЦСТ»*

Аннотация

Основной целью работы является разработка методики и исследование применимости нейросетей для тестирования графических интерфейсов в Linux системах. В работе предлагается новая композитная система тестирования и контроля качества программных продуктов. Одним из ключевых функциональных блоков системы является нейросеть, задача которой на первом этапе сводится к распознаванию элементов интерфейса и получению их координат на экране.

В работе обосновывается актуальность и значимость методики для нужд компании АО "МЦСТ" и индустрии в целом. Проведен анализ существующих автоматизированных средств тестирования и предлагается новая методика тестирования, описываются планы и возможные проблемы реализации. В дальнейшую реализацию проекта входят: изучение применимости нейронных сетей для использования их в тестировании, оценка надежности и эффективности новой методики, разработка прототипа готовой системы и проверка его на реальных задачах контроля качества компании АО "МЦСТ".

Введение и актуальность

Тестирование и контроль качества программных продуктов одно из ключевых направлений в современном мире. Количество программ самого разного назначения растет, увеличиваются темпы развития, и в разработку приходит все больше плохо квалифицированных кадров. При этом качество продукции и требования к ней только возрастают. Сейчас практически невозможно представить ни одну отрасль, где не использовалось бы программное обеспечение, а приложение, содержащие графический интерфейс пользователя, обязательно требование. АО "МЦСТ", как проектировщик микропроцессоров, разработчик и интегратор программного обеспечения на собственных аппаратных решениях, сталкивается с необходимостью такого тестирования постоянно. Из исключительно научного и промышленного решения МЦСТ все чаще требуется выпускать продукты, ориентированные для гражданского населения, а также современные переносные комплексы военного назначения. Эти новые для предприятия продукты неотделимы от графического интерфейса пользователя и программ, которые его содержат.

В настоящий момент в дистрибутивах семейства Linux, одним из которых является "ОС Эльбрус"[1], собственная разработка АО "МЦСТ", насчитываются несколько сотен программ, оснащенных графическим интерфейсом. Помимо очевидных затрат времени специалистов на ручное тестирование таких продуктов существует нетривиальная задача автоматизации такого тестирования, так как увеличение штата в современных реалиях практически невозможно по сугубо экономическим причинам. Автоматизация осложняется тем, что сообщество Linux пока не смогло выработать единого стандарта для доступа и контроля элементов интерфейса изнутри системы, что заставляет в ручном режиме прокликать приложения. При регулярном тестировании реального продукта это просто не представляется возможным. А при переносе программного обеспечения на аппаратную платформу e2k компании МЦСТ, о которой разработчик открытого ПО скорее всего никогда не слышал, становится критически важно.

Решением проблемы является использование систем автоматического тестирования интерфейсов, однако это осложняется сразу несколькими проблемами. Все имеющиеся системы имеют свои недостатки:

- Продукты заточены под ОС Windows;
- Проприетарные продукты;
- Продукты с ограниченным функционалом;
- Продукты, требующие заранее написанных сценариев.

Например, IBM Rational Functional Tester [2] заточен в большей степени под ОС Windows, является проприетарным и требует заранее написанные сценарии. Или же AutoBlackTest [3], который использует для своей работы IBM Rational Functional Tester.

Из-за наличия недостатков, которые довольно сильно ограничивают работу средств автоматического тестирования графических интерфейсов, появляется необходимость в создании собственного более универсального решения. В АО "МЦСТ" уже имеются отдельные небольшие системы, которые используют заранее записанные сценарии, они хорошо себя показали, но имеют критический недостаток – это необходимость постоянной поддержки сценариев в актуальном состоянии и отсутствия контроля промахов.

В качестве решения описанных проблем хочется разработать новую, композитную методику тестирования, одной из частей которой была бы нейронная сеть или набор нейронных сетей. Композитность системы обусловлена необходимостью ее дальнейшего расширения, чтобы использовать различные решения на разных этапах разработки, другими словами, разделить систему на блоки и дорабатывать каждый блок по отдельности, добиваясь высокой надежности, легкости, масштабирования и разработки, это распространенная практика. Нейросеть является одним из таких блоков, задача которой, на первом

этапе, отыскивать элементы интерфейса в знакомых и незнакомых для нее приложениях. В дальнейшем планируется использовать нейросеть для построения сценариев тестирования и для анализа корректности работы.

В настоящей работе ключевой целью является исследование возможностей методики и нейросетей для решения подобной нетривиальной задачи, а также разработка прототипа системы.

Методика тестирования графических приложений

По итогам разработки методики тестирования предполагается реализовать схему, примерная структура которой изображена на рис.1.

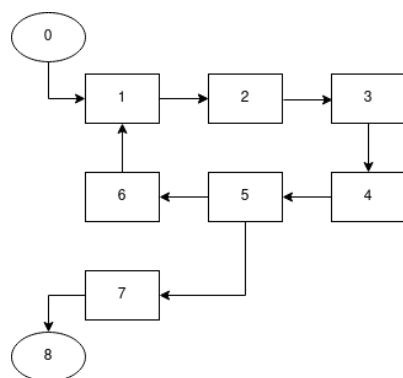


Рис. 1. Схема методики тестирования
(0 - вход, 1 - контроль состояния приложения,
2 - интерфейс получения изображения,
3 - анализатор изображения, 4 - составитель
сценариев, 5 - исполнитель сценариев, 6 - интерфейс
управления приложением, 7 - репортер, 8 - выход)

Схема методики тестирования с кратким указанием возможной реализации состоит из следующих элементов:

0 - инициация, старт тестирования;

1 - блок контроля состояния приложения. Анализирует состояние приложения и, как минимум, контролирует аварийное завершение работы через командную строку, но подразумевается, что он может быть и сложнее, в том числе, иметь свои нейросети, например, для анализа всплывающих окон и информации, содержащейся в нем;

2 - интерфейс получения изображения. Необходимо получить изображение текущего состояния приложения. Самым простым способом является printscreen экрана. Более сложным - анализ видеопотока камеры, направленной на экран тестируемой машины;

3 - анализатор изображения. Блок, целью которого является получение координат кнопок для последующей работы системы. Реализуется с помощью нейросети или набора нейросетей и является главной целью проводимого исследования;

4 - составитель сценариев. По полученным координатам кнопок данный модуль составляет сценарии тестов. Сценарий тестов - некоторая последовательность кликов мышью, ее перемещений

и нажатий клавиш. Простейшей реализацией является последовательность кликов;

5 - исполнитель сценариев. В его обязанности входит сравнение сценариев с известными, фильтрация сценариев и пр.;

6 - интерфейс управления приложением. В модуле происходит управление приложением. Реализация может быть различной. Для простоты можно использовать linux утилиту - xdotool [4], позволяющую имитировать работу компьютерных мыши и клавиатуры;

7 - репортер. Составляет отчет о прошедшем тестировании;

8 - конец тестирования.

При реализации методики, создании ее первого прототипа, главным становится исследование нейросетевых алгоритмов, задействованных в реализации блока 3. Именно этот блок является одним из новшеств настоящей работы. Все остальные блоки будут реализованы в самом простом виде или интегрированы уже из реализованных в МЦСТ систем. В перспективе, уже за рамками этого исследования они будут доработаны и оптимизированы.

План работы и ожидаемые проблемы

Для реализации поставленной задачи выделен следующий план, которого стоит придерживаться в исследовании алгоритмов и создании прототипа разрабатываемой системы.

Во-первых, необходимо провести обзор методов и способов тестирования различных приложений на различных ОС, включая обзор разработок в области тестирования черного ящика. С помощью кабинетного исследования необходимо найти разработки, которые могут нам помочь в настоящем исследовании.

Далее нужно дать обзор методам машинного обучения и провести исследование нейросетевых алгоритмов на возможность применения их в тестировании графических приложений, дать им оценку и определить, подойдут они нам или нужно создавать новую нейросеть.

Перед непосредственной разработкой алгоритма нужно составить датасет, в котором, как минимум, нужен скрин экранной формы и координаты кнопок. Данная задача очень рутинна, поэтому планируется реализовать алгоритм по быстрой генерации датасета – автоматическое создание интерфейсов.

На следующем этапе, в котором необходимо разработать методику, первичный вид которой представлен ранее, а также реализовать алгоритмы тестирования, сделав упор на создании модуля анализатора изображения.

Логичным продолжением предыдущего пункта плана является реализация прототипа, соединение всех реализованных модулей в одну систему, и его тестирование.

И, наконец, апробация результатов. Необходимо дать оценку работе в целом и реализованным алгоритмам, проанализировать их точность работы и сделать вывод об эффективности тестирования

графических интерфейсов с применением нейросетей.

При реализации такого проекта возможны критические трудности, которые необходимо иметь в виду уже на старте проекта. Ниже перечислены проблемы, часть из которых уже возникла, а с часть из которых потенциально возникнет:

- Неизвестно есть ли готовая нейросеть для решения задачи. Если есть, то необходимо понять какой нужно пользоваться, если нет, то нужно создавать свою, что усложняет задачу;

- Сложно оценить сколько времени уйдет на обучение нейросети и насколько быстро и эффективно она будет работать, что добавляет серьезный риск к исследованию;

- Непонятно как мы сможем ли мы добиться требуемой точности и как будем отслеживать промахи в сценариях;

- Не можем сказать сейчас, будет ли эффективен придуманный алгоритм автогенерации интерфейсов.

Для успешного завершения работы в срок необходимо придерживаться плана и при реализации иметь в виду перечисленные проблемы.

Заключение

В рамках работы была обоснована значимость тестирования приложений, имеющих графический интерфейс, выделены критические недостатки существующих систем, созданных для решения этой задачи. Предложена новая более универсальная методика тестирования, имеющая в своем составе нейросеть для анализа изображения программного окна. Намечены планы по реализации прототипа методики и выделены потенциальные сложности.

Сейчас создание прототипа находится в активной фазе исследования и разработки. За ближайшие полгода работу планируется завершить и получить работающий прототип, реализующий предложенную методику.

Актуальность работы в текущих мировых реалиях и ее важность и значимость для АО "МЦСТ" и других IT компаний не вызывает сомнений.

Список литературы

1. Операционные системы «Эльбрус». URL: http://www.mcst.ru/elbrus_os?ysclid=ldhohgsu8y678956439
2. IBM Rational Functional Tester. URL: <https://www.ibm.com/products/rational-functional-tester>
3. L. Mariani, M. Pezzè, O. Riganelli, M. Santoro. AutoBlackTest: Automatic BlackBox Testing of Interactive Applications. // 2012 IEEE Fifth International Conference on Software Testing, Verification and Validation
4. Xdotool - x11 automation tool. URL: <https://github.com/jordansissel/xdotool#readme>

ОБЗОР АНАЛИТИЧЕСКИХ МЕТОДОВ ОЦЕНКИ НАДЕЖНОСТИ ТЕХНИЧЕСКИХ СРЕДСТВ

Волков П.О.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии,
МИЭМ НИУ ВШЭ*

Аннотация

Данная работа представляет обзор распространенных методов расчета показателей безотказности, которые могут быть применены для технических средств разной степени структурной сложности монотонного типа. Это исследование необходимо, так как проводится в рамках разработки программного модуля для оценки надежности резервированных систем.

Введение

В эпоху глобальной автоматизации и постоянного усложнения структуры и функциональных возможностей вычислительных систем все более актуально становится задача повышения вероятности безотказной работы таких систем. Для этого проводится системный анализ надежности технических средств на всех стадиях жизненного цикла продукта с помощью специальных методов, в основном разработанных в период развития космической и военной отрасли в 1960х-1980х годах [1] и заложенных в соответствующих государственных стандартах.

Ввиду большого разнообразия методологий анализа надежности вычислительных средств данная работа представляет краткий обзор классических методов оценки параметров безотказности монотонных систем со сложной структурой: вероятности безотказной работы, коэффициента готовности. Методы можно условно разделить на простые и высокой степени сложности в зависимости от сложности анализа (построения математической модели) рассматриваемых структурных схем надежности.

Простые методы оценки надежности

Для эффективного анализа и расчета показателей надежности часто используются методы структурного расчета [2]. Они заключаются в представлении объекта в виде структурно-логической схемы, описывающей в общем случае логические связи между состояниями элементов и объекта в целом, с учетом структурно-функциональных связей и взаимодействия элементов. Данная топологическая модель в ряде случаев может также учитывать стратегии обслуживания, виды и способы резервирования. Особой популярностью обладает такая структурная схема, как блок-схема надежности (RBD), представляющая техническое средство в виде набора блоков, связанных согласно политике надежности данной системы [3]. В зависимости от желаемого уровня разукрупнения устройства, блоком может считаться как отдельный радиоэлемент, так и группа радио-

элементов. RBD дает достаточно наглядное представление о надежности данного объекта, особенно по его возможным последовательным и параллельным соединениям. Стоит отметить, что большинство методов, связанных с построением и анализом RBD, ограничиваются только такими вычислительными системами, где состояние отдельного блока может находиться только в одном из двух состояний: работоспособным или нет, а отказ одного блока не влияет на работоспособность другого; иначе нужно использовать другие способы оценки надежности.

Также стоит обратить внимание, что в общем случае из-за сложной структуры построенных технических систем, перед непосредственным анализом RBD и дальнейшим выводом оценок надежности проводят редукцию исходной большой RBD, получив несколько (малых) суб-RBD [4]. Единственное ограничение данного разложения является то, что его очень трудно автоматизировать.

Простейшими подходами к RBD-анализу можно считать использование теоремы о полной вероятности или использование таблицы истинности, которые рассмотрены в соответствующих государственных стандартах [4]. Первый метод подразумевает использования взаимоисключающих событий, которыми, как правило, являются работоспособность и неработоспособность какого-то блока. В простейшем случае формула может иметь вид (1):

$$P_s = P(S|X_p) * P(X_p) + P(S|X_{не_p}) * P(X_{не_p}) \quad (1)$$

Метод таблицы истинности определяет работоспособность всех возможных состояний системы при различных сочетаниях работоспособных и неработоспособных элементов. Хотя таблица истинности может быть уменьшена булевыми методами, ее может быть трудно применить к сложной системе, поскольку количество состояний может быстро стать очень большим.

Такие методы могут быть полезны для довольно небольших RBD, которые изначально не приведены к последовательно-параллельному типу, такие как мостиковые диаграммы (рис.1).

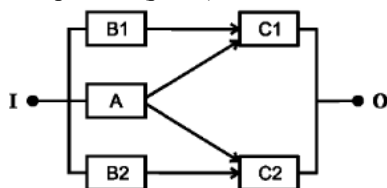


Рис. 1. RBD на примере мостиковой схемы

Существуют также частные случаи построения и работы технического средства, для которого проще вывести отдельные формулы, а не составлять каждый раз таблицу истинности. Например, это может быть мажоритарная система m из n , которая устанавливает работоспособность системы только при том условии, что работоспособными являются не менее m элементов из n [4]. Такие системы часто используются при резервировании. Так, вероятность безот-

казной работы этой системы (но идентичных элементов) может быть вычислена методом производящих функций по формуле (2), где используется число сочетаний C .

$$P_s = \sum_{k=m}^n C_n^k * P^k * (1-P)^{n-k} \quad (2)$$

Вместо составления таблиц истинности при анализе RBD может применяться метод булевой модели [4], то есть математической модели, в которой состояние системы представляется логической функцией булевых переменных, представляющих состояния компонентов системы в 1 или 0. Логическое (булево) выражение успеха системы для сложной структуры может иметь зависимые события работоспособного состояния блоков (пересекающихся наборы блоков [3]). Поэтому разрабатываются специальные алгоритмы формирования непересекающихся наборов, которые позволяют заменить булевы переменные на соответствующие показатели надежности. Например, это может быть метод дизъюнкции [4] или отдельное исследование [9], где авторы статьи широко используют постулаты, соответствующие алгебре логики, для составления логической функции работоспособного состояния системы. В некоторых случаях построение таких функций эффективнее и нагляднее, чем построение RBD.

Методы оценки надежности высокой степени сложности

Более совершенным методом оценки показателей надежности по RBD, но чуть трудоемким является метод минимальных путей и минимальных сечений [3]. Соответствующий стандарт устанавливает, что если RBD рассматривать как электрическую цепь, то минимальный путь - это минимальный замкнутый контур, ведущий к работоспособному состоянию цепи, а минимальное сечение соответствует минимальному размыканию цепи, ведущий к нерабочему состоянию. Благодаря относительно простым алгоритмам, таким как карты Карно или бинарная диаграмма решений (BDD), из RBD можно извлечь минимальные пути и сечения [3]. Наиболее часто используемый метод - BDD - компактное дерево принятия решений (рис.2), основанное на разложении Шеннона логического выражения.

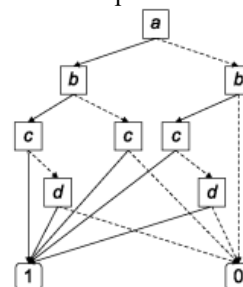


Рис. 2. Пример BDD

Так как минимальные пути и сечения представляют собой пересекающиеся наборы блоков, для определения, например, вероятности безотказной работы обычно используется формула Сильвестра-

Пуанкаре. Например, для минимальных путей формула может иметь вид (3):

$$P_s = \sum_i P(\Pi_i) - \sum_{i < j} P(\Pi_i * \Pi_j) + \sum_{i < j < k} P(\Pi_i * \Pi_j * \Pi_k) - \dots \quad (3)$$

Стоит отметить, что из-за громоздкости формулы и сложности поиска всех путей эту формулу при определенных условиях можно аппроксимировать на первом члене. Помимо прочего, часто бывает более оптимально и предпочтительно найти не точные значения оценок надежности, а интервальные, как предлагает Ушаков И.А. [5].

Другой распространенной формой структурной схемы надежности может быть диаграмма состояний и переходов, описывающая возможные состояния объекта и его переходы из одного состояния в другое (рис.3).

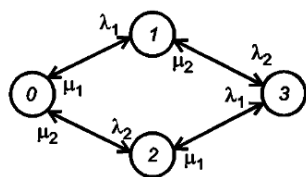


Рис. 3. Пример диаграммы состояний и переходов

Причем необходимо обеспечить, чтобы эта диаграмма обладала тем свойством, что будущее состояние системы зависело только от текущего состояния системы, а не от того, как система попала в это состояние. Для анализа таких диаграмм используются методы марковского анализа [6]. Решение марковских моделей надежности традиционными методами сводится к составлению системы однородных дифференциальных уравнений, которые описывают поведение исследуемой системы и решение которых приведет к нахождению требуемых параметров надежности [5]. Основное преимущество применения методов марковского анализа заключается в том, что стратегии обслуживания могут быть легко смоделированы. Кроме того, модель может отражать порядок возникновения множественных отказов, а также она полезна для систем, в котором число возможных состояний элементов системы больше двух. Однако этот метод имеет некоторые ограничения. Во-первых, с увеличением количества компонентов системы количество состояний увеличивается экспоненциально. Во-вторых, пользователям может быть сложно построение модели марковского процесса и управление ею. Среди прочего, решение системы дифференциальных уравнений может быть трудоемкой задачей, хотя и есть исследования в этой области [7], в которых Шубинский И.Б. и Замышляев А.М. предлагают топологический полумарковский метод для вычисления параметров надежности технических систем согласно строгим формульным выражениям.

Из-за многообразия технических средств вполне могут быть, например, системы с зависимыми (коррелированными) элементами, то есть когда отказ одного блока может привести к отказу другого (не идентичного). Создание методов анализа надежности таких систем является вполне сложной, но решаемой

задачей. Например, такой задачей занимались С. Парк, С. Чой, К. Сикорский, Н. Стаббс в соответствующей статье [8]. К тем же сложным (в смысле анализа надежности) системам могут быть устройства с зависимыми от времени параметрами элементов.

Заключение

В данной работе был представлен ряд основных практико-ориентированных методов оценки надежности сложных технических систем произвольной монотонной структуры. Каждый из методов обладает своими преимуществами и недостатками (ограничениями), связанными с подходами к анализу технических средств, а также со свойствами элементов, составляющих эти системы. Поэтому нет единого оптимального метода по оценке надежности любого технического средства. Следовательно, можно заключить, что наилучшим способом решения данной задачи является комбинирование нескольких методов, применение каждого из которых зависит от свойств устройства и желаемых условий оптимальной реализации на ЭВМ. Так, если система последовательно-параллельного типа (связанного с резервированием и другими особенностями) обладает теми свойствами, что каждый элемент ССН может находиться в одном из двух состояний – работоспособность или отказ, а отказы элементов ССН – события независимые, то можно воспользоваться простыми булевыми и вероятностными методами. Если система или ее подсистема более сложной структуры (мостиковые и другие схемы), то рационально попробовать применить метод минимальных путей и сечений. И, наконец, если система требует сложной стратегии технического обслуживания или ее элементы могут находиться в более чем двух состояниях, обычно применяют марковский анализ построенных диаграмм состояний и переходов (на основе решения системы дифференциальных уравнений).

Список литературы

1. Основы теории надежности - Тимошенков С.П.
2. ГОСТ 27.301-95. Надежность в технике. Расчет надежности. Основные положения
3. ГОСТ Р МЭК 61078-2021. Надежность в технике. Структурная схема надежности
4. ГОСТ Р 51901.14-2007. Структурная схема надежности и булевы методы
5. Ушаков И.А. (1985) Надежность технических систем: Справочник
6. ГОСТ Р 51901.15-2005. Применение марковских методов
7. Shubinsky Igor B., Zamyshlyayev Alexey M. - Topological semi-Markov method for calculation of stationary parameters of reliability and functional safety of technical systems
8. S. Park, S. Choi, C. Sikorsky, N. Stubbs - Efficient method for calculation of system reliability of a complex structure.
9. Dilshod Okhunov, Mamatjon Okhunov, Mukaddas Akbarova - Method of calculation of system reliability on the basis of construction of the logical function of the system.

РАЗРАБОТКА АЛГОРИТМА ДЛЯ СОЗДАНИЯ ДАТАСЕТА НА РУССКОМ ЯЗЫКЕ ДЛЯ ЗАДАЧИ ПРЕОБРАЗОВАНИЯ ТЕКСТА В РЕЧЬ

*Смоленчук И.К., Нижегородов М.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Датасет представляет собой набор аудио отрезков, длиной около 10 секунд для более быстрой обработки, и их транскрипция, выделяемая при помощи алгоритма с Speech-to-text моделью. Эти данные помещены в csv файл для дальнейшего использования и обучения моделей преобразования текста в речь. Общая продолжительность непрерывной речи около 37 часов.

Введение

В данный момент задачи, связанные с обработкой естественного языка (NLP), становятся все популярнее и мы часто встречаемся с ними в повседневной жизни, например, голосовые помощники, переводчики и т.д. Но для этих задач необходим большой объем данных, таких как речь и ее письменное представление. Например, известный датасет RUSLAN [1], имеет более 31 час непрерывной речи или англоязычный LJ Speech на 24 часа непрерывной речи [2].

Сейчас в открытом доступе нет большого числа русскоязычных датасетов, которые могут выполнять задачи по тренировке text-to-speech моделей, в отличие от англоязычных. Процесс создания таких датасетов может включать большое количество трудностей и быть довольно трудоемким, ведь записей голоса недостаточно, еще необходимо разделить запись на короткие отрезки, которые разделяются паузами в речи, чтобы дальнейшая модель смогла использовать такой большой объем данных. Также необходимо убрать артефакты в начале и конце отрезков, где слово может быть обрезано, либо начинаться слишком резко. Транскрипция речи может занять долгое время, поэтому при разработке датасета в алгоритм была вставлена Speech-to-text модель, которая размечает отрезки аудио и на выходе будет датасет, состоящий из отрезков и их разметки.

Целью данной работы является создание полноценного русскоязычного датасета, поскольку проблема с объемом качественных данных в данной сфере наблюдается. В ходе анализа поставленной задачи, была выявлена необходимость создания полноценного алгоритма, который бы смог не просто создавать датасет – объединять в csv файл набор аудиозаписей с транскрипцией, но и являлся комплексным решением, которое было бы удобно конечному пользователю, а именно упростить процесс сбора наборов данных. Исходя из этого, появилось необходимость не просто кластеризовать аудиозаписи, а также создавать аудиофайлы на основе большой записи, но и делать это без потери качественных показателей, например, избегать разбиения

аудио на участки в местах, где диктор говорит фразу, в результате чего, в таких местах будет наблюдаться разрыв целостности аудиофайлов.

Поэтому комплексная цель данной работы есть создания датасета и алгоритма преобразования аудио в отрезки необходимой длины, который также дополняет и улучшает качество этих отрезков, а также описывает процесс оценки STT модели, которая используется для генерации разметки.

Сбор данных для аудио коллекции

Для создания качественного датасета, позволяющего использовать данные для обучения TTS моделей с хорошей точностью, необходимо собрать большое количество аудио данных. В данной работе за источник этих данных была взята электронная библиотека выпусков [3] от известного видео блогера.

Были взяты выпуски, включающие речь диктора в формате монолога и находящиеся в открытом доступе, и отредактированы для исключения лишнего шума и посторонних голосов (путем удаления таких фрагментов). Общая продолжительность аудиофайлов составила 37 часов непрерывной речи диктора.

Алгоритм семплирования и кластеризации звуковых данных

STT модель является довольно объемной и сложной структурой, выбранная в данной работе модель, имела около миллиарда параметров. Исходя из этого, для работы качественной Speech to text модели необходимо крайне много ресурсов, которые увеличиваются с каждой дополнительной секундой длины входного аудио файла, именно для снижения нагрузки на вычислительный модуль и был разработан алгоритм семплирования аудио, суть которого заключается в разбиении большого аудио файла, для обработки которого, допустим потребовалось бы 100 GB видеопамати на 100 небольших сэмплов длиной значительно короче исходного аудиофайла, для обработки каждого, при этом, потребуется лишь 1 GB видеопамати. Данный алгоритм является крайне гибким, поскольку может получать данные в разнообразных форматах. На текущий момент проработаны варианты работы в форматах .wav .mp3 путем использования функции конвертирования из одного формата в другой. Основу алгоритма составляет модуль Pydub для языка Python [4]. На основе методов обработки звука реализован способ семплирования аудиозаписей на отдельные “кусочки”. Разбиение реализовано при помощи оценки пауз диктора, основанный на поиске заданных параметров тишины, а именно поиск отрезка в аудио, где звуки не превышают определенного настраиваемого порога, например -16DBFS [5], означает, что если алгоритм находит участок, который тише данного значения, данный участок может являться местом разбиения сэмплов. Второй настраиваемый параметр - длина выбранного участка, данный параметр позволяет сделать алгоритм более гибким, а именно разбивать речь человека не на предложения а, например, на смысловые участки - места где каждый человек под-

сознательно ставит запятую и набирает воздуха, то есть делает небольшую паузу. Наконец, выделенные сэмплы проходят этап проверки длины, чтобы не получить набор аудиозаписей длиной по 1 секунде, выбирается порог длительности, например - 5 секунд, что означает что несколько коротких участков будут записываться в один сэмпл, пока их длина не станет больше чем 5 секунд. Данный параметр также легко настраивается. Затем, для улучшения качества распознавания речи, в начало и конец каждого интервала вставляются пробелы, чтобы ASR модель лучше распознала начальные и конечные слова. После выполнения вышеописанного алгоритма на собранных аудиозаписях собственная коллекция включает 22889 отрезков продолжительностью от 5 до 10 секунд. Благодаря распознаванию речи, на выходе получается готовая разметка, с некоторыми небольшими промахами, которая проверяется алгоритмом на орфографию (так как речь распознается по морфемам, а не словам целиком). Корректированная разметка сопоставляется с названием своего отрезка аудио и записывается в конечный csv-файл, структура которого состоит из 2 элементов: название аудио отрезка и его разметка. Этот файл, в совокупности с файлами сэмплов, представляют собой готовый датасет для задач синтеза речи.

Для улучшения качества работы алгоритма были определены выделенные гиперпараметры эмпирическим путем. Например, при выборе слишком высокого порога шума, алгоритм просто не мог выделять отдельные части предложения и создавал слишком большие аудио сегменты, которые не подходили для задания, а при выборе слишком короткой продолжительности тишины наблюдалась обратная ситуация, которая бы, в целом могла бы теоретически быть приемлемой, однако при дальнейшем прослушивании возникали артефакты – при многократной нарезке небольшого аудио фрагмента – допустим 5 секунд, например, на 50 промежутков, и повторной их склейке в 1 аудио фрагмент при последующем прослушивании возникает четкая проблема, слышно, что человек говорит обрывисто, поскольку минимальные потери при нарезке аудиозаписи все же существуют, из за этого, в дальнейшем, при слишком высокой дифференциации аудио, обучаемая модель могла бы вести себя непредсказуемо, поскольку она теряла бы части слов. Также, добавление 1.5 секунды тишины в конец каждого аудиосегмента было тоже выведено эмпирически вследствие того, что модель захватывает не всю аудиозапись для распознавания, например, как при голосовых сообщениях в популярных мессенджерах, часто не слышно последнего сказанного слова, при транскрипции аудиозаписей, была установлена необходимость добавления такой доработки, которая позволила работать ASR модели более стабильно.

Методология выбора STT модели

В алгоритме используется модель преобразования речи в текст, для разметки полученных отрезков речи. Для хорошего качества датасета необходима хорошая разметка, но без ручной. Поэтому в алгоритме

необходима эта модель - она исключает необходимость вмешиваться в работу алгоритма, который самостоятельно нарезает аудио и получает их транскрипцию, ведь при ручной разметке всего аудио затем будет необходимо сопоставлять нарезанные аудио и их текст.

Для получения качественной разметки были взяты несколько STT моделей, которые находятся в открытом доступе и проверены на тестовом датасете, имеющем эталонную разметку. Для оценки точности распознавания речи использовался процент неправильно распознанных слов WER - Word Error Rate [6]. Метод определения WER состоит в выравнивании двух текстовых строк, где первая — это результат распознавания, а вторая - эталонная разметка), с помощью алгоритма расчета расстояния Левенштейна. Это расстояние представляет собой “стоимость” редактирования данных для преобразования первой строки во вторую с наименьшим числом операций ручной замены (S), удаления (D) и вставки (I) слов:

$$WER = \frac{S+D+I}{T}, \quad (1)$$

где T - количество слов в распознанной строке.

Для оценки качества существующих моделей Automatic Speech Recognition были взяты 3 модели – Google Speech Recognition [7], Wav2Vec [8] и Whisper [9]. Проверка осуществлялась при помощи тестового датасета, который состоит из 64 аудио отрезков, общей продолжительностью около 6 минут, собранных из полного аудио файла, находящегося в общем датасете. Данные отрезки были размечены вручную для получения стопроцентной точности разметки. Алгоритм оценки включал в себя сравнение предсказанной и эталонной разметок, вычисление WER и CER метрик для каждого отрезка и вычисление средних показателей этих метрик для всех 64 отрезков.

После вычисления были получены следующие результаты: модель Wav2Vec являлась наименее точной с процентом ошибочно распознанных слов – 27,34% и ошибочно распознанных символов – 10,57%, а модели GSR и Whisper показали довольно равные результаты, так как у модели от Google значения ошибочно распознанных слов и символов – 12,9% и 3,5% соответственно, при показателях Whisper в 12,6% и 4,9% соответственно. Из полученных показателей был сделан выбор в пользу модели от компании Google за счет ее меньшей ошибки в распознанных символах, поскольку модель делает ошибки в словах немного чаще, но при этом кол-во этих ошибок в словах на 1,5% меньше.

Заключение

Таким образом, описанный алгоритм упростит решения по созданию датасетов путем автоматизации процесса разбиения аудиозаписей на отрезки, процесса преобразования этих отрезков к модели STT, а также автоматизации разметки этих отрезков и приведение данных к нужному виду. Помимо этого, созданный при помощи алгоритма датасет позво-

лит расширить объем общедоступных данных для обучения NLP моделей, позволяющих синтезировать речь из текста голосом определенного человека, чья речь была подана на вход самого алгоритма.

Список литературы

1. Gabdrakhmanov L., Garaev R., Razinkov E. (2019) RUSLAN: Russian Spoken Language Corpus for Speech Synthesis. In: Salah A., Karpov A., Potapova R. (eds) Speech and Computer. SPECOM 2019. Lecture Notes in Computer Science, vol 11658. Springer, Cham
2. The LJ Speech: [сайт]. – 2017 – URL: <https://keithito.com/LJ-Speech-Dataset/> (дата обращения 23.01.2023). – Текст : электронный
3. Oper.ru: [сайт]. - 1999 - URL: <https://oper.ru/> (дата обращения: 25.01.2023). - Текст : электронный
4. Github: [сайт]. - 2008 - URL: <https://github.com/jiaaro/pydub> (дата обращения: 25.01.2023). - Текст : электронный
5. Wikipedia: [сайт]. - 2001 - URL: <https://en.wikipedia.org/wiki/DBFS> - (дата обращения: 25.01.2023). - Текст : электронный.
6. Карпов А.А., Кипяткова И.С. Методология оценивания работы систем автоматического распознавания речи // Приборостроение. 2012. №11. URL: <https://cyberleninka.ru/article/n/metodologiya-otsenivaniya-raboty-sistem-avtomaticheskogo-raspoznavaniya-rechi> (дата обращения: 24.01.2023).
7. Google Cloud: [сайт]. - 2007 - URL: <https://cloud.google.com/speech-to-text> - (дата обращения: 25.01.2023). - Текст : электронный.
8. wav2vec2-large-xlsr-53 Hugging Face [сайт]. URL: <https://huggingface.co/facebook/wav2vec2-large-xlsr-53> (дата обращения: 25.01.2023).
9. whisper-large Hugging Face [сайт]. URL: <https://huggingface.co/openai/whisper-large> (дата обращения: 25.01.2023).

ФОРМАЛИЗАЦИЯ МЕТОДИКИ ОЦЕНИВАНИЯ СОСТОЯНИЯ ПРОГРАММНЫХ ПРОЕКТОВ ПО КРИТЕРИЯМ ПРОГРЕССА И ЗДОРОВЬЯ ПРОЕКТА НА БАЗЕ СТАНДАРТА OMG ESSENCE

Гаранина А.Е., Иванова Е.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматривается модель, построенная на основе стандарта Essence, с помощью которой возможно отслеживание качества и “здоровья” проектных работ, выполняемых в рамках учебных заведений. Формализованная модель, может лежать в основе автоматизации получения оценки результатов проектной деятельности. С ее помощью проанализированы студенческие проекты и было подтверждено,

что модель в совокупности с деревом взаимосвязей между основными показателями проекта может предполагать причины неуспеха проекта.

Введение

Ежегодно студенты высших учебных заведений участвуют в разработке и сопровождении иногда нескольких проектов, которые проходят в рамках учебной и дополнительной программы. В связи с этим важно понимать, как члены команды взаимодействуют между собой, разбираются в предметной области и борются с возникающими проблемами.

Если существует потребность в проведении проектных работ, то существуют и сопутствующие проблемы с отслеживанием хода этих работ. Часто члены команды могут быть незнакомы до начала проекта, плохо коммуницировать в группе, а иногда и вовсе не понимать сути и хода проекта. После чего в проектной работе могут возникать трудности, отступление от сроков, недопонимание между командой и заказчиком, и в конечном счете - неудовлетворительный результат всего проекта, когда ожидания совсем не совпадают с действительностью.

Исходя из вышеописанного выявляется проблема, состоящая в необходимости в ходе проекта оценивать не то, что участники работают, а то, каких результатов они добиваются своим трудом. В связи с чем можно определить задачу получения инструмента, который даст возможность автоматизировано выявлять проблемные места в ходе проведения проектных работ.

Анализ различных методов и методологий оценки качества [1-6] показал, что стандарт OMG Essence [7, 8], разработанный международным сообществом SEMAT [7-10] под руководством Ивара Якобсона [11, 12], в наибольшей степени подходит для наших нужд.

В связи с актуальностью данной проблемы была поставлена задача исследования возможности адаптации к автоматизации методики выявления проблемных мест ведения оценки проектных работ в соответствии со стандартом Essence, разработанного в рамках единой теории программной инженерии Semat Object Management Group (OMG).

Согласно Essence, оцениваются прогресс (увеличение количества продукта проекта) и “здоровье” проекта, то есть то, насколько эффективно с позиций достижения правильных промежуточных результатов развивается проект в заданных точках его развития.

В рамках задачи необходимо было провести тестирование методики на ряде проектных работ и провести ее адаптацию к возможности автоматизации получения формальной оценки нарушений «правильных» состояний конкретными участниками проекта.

Метод оценки качества и “здоровья” проекта

Ядро Essence определяет семь сущностей (по-другому их называют “альфами”), которые являются общими для всех проектов и которые взаимодействуют между собой: заинтересованные стороны,

возможность, требования, программная система, работа, команда, технология работы [13, 14, 15]. Для этих сущностей в рамках МИЭМ были адаптированы под студенческие проекты чек-листы, описанные в рамках стандарта [13].

Для автоматизации получения оценки качества и “здоровья” проектных работ описана математическая модель [14], построенная на основе предикатов первого порядка и чек-листов для всех циклов реализации проекта, которая дает возможность автоматизировать выявление причин неуспеха и выработку рекомендаций по оценке состояния проекта.

Значения предикатов первого порядка определяются из чек-листов, то есть из результатов, достигнутых в каждом цикле.

Если альфа на данном этапе достигла необходимого состояния, то оно маркируется единицей, иначе – нулем. По чек-листам можно определить сколько единиц должно быть на каждом этапе для достижения нормального состояния.

Рассмотрим на примере. На первом цикле сущность “заинтересованные стороны” должна принимать значение “признаны”, для этого должен выполняться список условий [13]:

- Выявлены все различные группы заинтересованных сторон, которые затронуты или будут затронуты разработкой и эксплуатацией системы программного обеспечения.
- Существует соглашение о группах заинтересованных сторон, которые будут представлены.
- Как минимум, заинтересованные группы, которые финансируют, используют, поддерживают систему, рассмотрены.
- Определены обязанности представителей заинтересованных сторон.

Состояние “признаны” будет достигнуто (примет значение 1) только в том случае, если чек-лист будет иметь значения (1,1,1,1), то есть все условия из списка выполнены, иначе - 0.

Такое предположение можно представить в виде предикатов, начиная с чек-листов, и заканчивая итогами по всем циклам для всех альфа.

Для состояния “признаны” функция будет выглядеть следующим образом:

$$\varphi_{\text{признаны}} = \begin{cases} 1, & \text{если } P(1,1,1,1) \\ 0, & \text{иначе} \end{cases}$$

Аналогичным образом можно описать состояния каждой сущности на каждом цикле, таких циклов 8.

Таким образом, становится понятно, что легко выявить причины того, почему состояние аномальное и сущности, на которые следует подействовать. Действительно, если значение предиката равно 0, то выбираем для соответствующей альфы те состояния, которые были равны 0, для них вероятнее всего не выполнены условия (достижение правильных состояний), что потенциально и указывает на то, что является причиной аномального состояния. Среди всех кандидатов на неправильные состояния выявляем, связывают ли их отношения (уровня ядра). Если да, то неправильное состояние может быть не из-за са-

мого состояния, в котором обнаружена аномалия, а из-за того состояния, которое является участником отношения с рассматриваемым состоянием.

Для рассмотрения такого возможного влияния было составлено дерево взаимосвязи сущностей через отношения по ядру ESSENCE (kernel), на котором представлены отношения сущностей друг с другом. Дерево представлено на рис. 1.



Рис. 1. Деревья взаимосвязи сущностей через отношения

Стоит отметить, что левый граф раскладывается в дерево и с точки зрения Vee-модели левый граф соответствует левой ветви модели (разработка), а правый – правой ветви модели (проверка).

Таким образом, можно составить алгоритм исследования качества и здоровья проекта:

- выявление нормального состояния на данном этапе для каждого альфа;
- оценка текущего состояния (составление чек-листов);
- на основании предикатной модели определение сущностей, которые не достигли нормального состояния;
- по дереву взаимосвязи определение сущностей, которые участвовали в получении данного состояния (определяются причины);
- по графику определение дисбаланса между нормальным и полученным состоянием (определяется здоровье проекта);
- составление рекомендаций по улучшению состояния проекта.

Тестирование

Для проверки разработанной математической модели были проведены эксперименты на нескольких проектах, которые проходили в рамках проектной деятельности в МИЭМ НИУ ВШЭ.

Руководители или тим-лидеры команд заполняли чек-листы на каждом цикле проектной работы. Полученные таким образом данные были проанализированы, по заполненным чек-листам были построены графики, в которых отражается нормальное состояние проекта, а также состояние, которое принимают сущности в исследуемом проекте. Далее с помощью дерева взаимосвязей между альфами определялась причина недостижения нормального состояния на исследуемом цикле.

Рассмотрим результаты проведенного тестирования на примере одного из проектов. На рис. 2 отображены результаты для сущности “заинтересованные стороны”.

Заинтересованные стороны	
Получены	0
0	Выявлены все различные группы заинтересованных сторон, которые затронуты или будут затронуты разработкой и эксплуатацией системы программного обеспечения.
0	Существует соглашение о группах заинтересованных сторон, которые будут представлены.
0	Названы группы, заинтересованные группы, которые финансируют, используют, поддерживают систему, рассмотрены.
1	Определены обязанности представителей заинтересованных сторон.
Представлены	0
0	Представители заинтересованных сторон согласились взять на себя свои обязанности.
0	Представители заинтересованных сторон уполномочены выполнять свои обязанности.
0	Планы взаимодействия между представителями заинтересованных сторон были согласованы.
1	Представители заинтересованных сторон поддерживают и уважают способ работы команды.

Рис. 2. Результаты чек-листа на первом цикле сущности “заинтересованные стороны” одной из проектных команд

Как видно на рисунке, не все альфы заполнены. В данном случае математическая запись будет выглядеть следующим образом:

$$\varphi_{\text{признаны}} = 0, \text{ т. к. } P(0,0,0,1)$$

Аналогично,

$$\varphi_{\text{представлены}} = 0, \text{ т. к. } P(0,0,0,1)$$

Следовательно,

$$\varphi_{\text{Заинтересованные стороны}} = 0, \text{ т. к. } P(0,0).$$

В целом по циклу

$$\varphi_{\text{цикл 1}} = 0, \text{ т. к. } P(0,0,0,0,0,0,0).$$

Нормальная математическая модель по данному циклу выглядит так:

$$\varphi_{\text{цикл 1}} = 1, \text{ при } P(1,1,1,0,0,1,0).$$

На данном цикле проект не достиг нормального состояния. Далее строится график (см. рис. 3) и проверяется взаимосвязь альф по дереву связи для того, чтобы определить слабые места проекта на данном цикле.

На графике синим изображено принимаемое значение, красным - нормальное. Для определения возможных причин ненормальных состояний применим для этого графика дерево взаимосвязей. Можно заметить, что только альфа “Команда” приняла нормальное состояние, все остальные - нулевое. Так как это учебный проект, это может быть связано с тем, что команда только сформировалась, но не были представлены заинтересованные стороны, не обозначены возможности и требования.



Рис. 3. Нормальные и принимаемые значения сущностей на первом цикле проекта

То же самое было проделано с каждой альфой на каждом цикле проектной работы и в конце анализа наблюдается взаимосвязь, которая прослеживается между альфами, с помощью которой отслеживаются

слабые места проекта и причины, которые влияют на неудачу цикла.

Заключение

В ходе проведенного исследования была проведена адаптация стандарта OMG Essence к возможности автоматизации получения формальной оценки нарушений «правильных» состояний участниками проекта и проведено тестирование на ряде проектов. Это стало возможно за счет формального описания методики в виде математической модели и описания взаимосвязей между сущностями в виде дерева решений.

Таким образом, развитие методики было достигнуто за счет конкретизации Essence применительно к общей методике оценки проектов в МИЭМ, чтобы можно было видеть развитие однотипных проектов.

Полученные результаты в дальнейшем могут быть оформлены в виде приложения, которое при проставлении чек-листов автоматически будет выявлять причины “неуспеха” в проекте и демонстрировать рекомендации пользователю.

Полученные результаты могут быть использованы при оценке проектных работ в учебных заведениях.

Список литературы

1. P. Kruchten. A conceptual model of software development. Software Project Management with OpenUp, 2007.
2. W.S.Humphrey. Why Teams Need Operational Processes. Software Engineering Institute, Carnegie Mellon University, 2009.
3. B. Boehm, A.Jain. An initial theory of value-based software engineering. Value-based Software Engineering, Springer, 2005.
4. Capers Jones. Software quality in 2010: A survey of the state of the art. Capers Jones & Associates, 2010.
5. D. Bjorner. The Triptych process model – process assessment and improvement. Tokio University, 2006.
6. Capers Jones. Analyzing the tools of software engineering. Capers Jones & Associates, 2010.
7. ESSENCE User Guide // SEMAT URL: <https://semat.org/view-1-essence-lite> (дата обращения: 08.02.2022).
8. Essence – Kernel and Language for Software Engineering Methods: официальный сайт. - URL: <https://www.omg.org/spec/Essence/1.2/PDF> (дата обращения: 10.02.2022)
9. Semat : официальный сайт. - URL: <https://www.semat.org/> (дата обращения: 05.02.2022).
10. OMG Essence : официальный сайт. - URL: http://sewiki.ru/OMG_Essence (дата обращения: 08.02.2022).
11. “The Essentials of Modern Software Engineering.” / Ivar Jacobson, Harold “Bud” Lawson, Pan-Wei Ng, Paul E. McMahon, Michael Goedicke, 2019.
12. “The Essence of Software Engineering” Applying the SEMAT Kernel”/ Ivar Jacobson, Pan-Wei Ng, Paul E. McMahon, Ian Spence, Svante Lidman. Addison – Wesley, 2013.
13. Проектное обучение студентов: методическое пособие / д.т.н., профессор Позин Б.А., Зими́на Е.Ю.,

к.э.н. Кирюшина А.А., Ковальчук И.М., Кочнев А.А., Новиков Р.С., Свинцицкая Е.Б., к.т.н., профессор Старых В.А. Москва, 2020.

14. Змеев Д. О. Прототип системы поддержки принятия решений для управления проектами на основе стандарта OMG ESSENCE и байесовских сетей: диссертация. Томск, 2022.

15. Позин Б. А., Горбунова Е. Развитие базовой модели Semat для жизненного цикла заказных ответственных программных систем // В кн.: Материалы четвертой Научно-практической конференции "Актуальные проблемы системной и программной инженерии". Сборник трудов. М.: НИУ ВШЭ, 2015.

РАЗРАБОТКА СИСТЕМЫ ПРЕДОТВРАЩЕНИЯ АВАРИЙНЫХ СИТУАЦИЙ ПРИ 3D ПЕЧАТИ

*Горбачева В.Ю.¹, Жукова М.А.²,
Зарубина Т.В.³, Ильинская Е.В.³, Палуха Д.В.¹*
Национальный исследовательский университет

«Высшая школа экономики»,

¹департамент компьютерной инженерии,

²департамент электронной инженерии,

³департамент прикладной математики

МИЭМ НИУ ВШЭ

Аннотация

В работе рассматриваются возможные классификации ошибок при 3D-печати, создание размеченного набора данных, а также разработка программного обеспечения, способного выявлять аварийные ситуации в режиме реального времени в процессе печати на 3D-принтере и уведомлять об этом оператора.

Введение

Технологии 3D-печати развиваются чрезвычайно быстро и используются практически во всех сферах жизни. Во время печати на 3D-принтере могут возникать такие дефекты как обрыв нити, отрыв детали от стола, засор экструдера и прочие нештатные ситуации. Несвоевременное их выявление может привести к порче материалов и выходу из строя оборудования, но дополнительное программное обеспечение поможет оператору вовремя устранить неполадки и избежать серьезных последствий.

Таким образом, необходимость в разработке программного обеспечения (ПО) для предотвращения аварийных ситуаций при работе с 3D принтером является актуальной. Разрабатываемое ПО использует обработку изображений с портативных камер, установленных на принтере, для отслеживания ошибок при печати и датчики, сигнализирующих о них.

Создание размеченного набора данных

Разрабатываемое программное обеспечение включает в себя взаимодействие с нейронной сетью, которая способна распознать аварийную ситуацию по изображению с камер. Для её обучения необходимо создать размеченный набор данных, состоящий

из фотографий аварийных ситуаций и их классификации.

При обзоре возможных ошибок при 3D-печати, для обучения нейронной сети были выделены 16 категорий [1] аварийных ситуаций. В каждой из категорий находятся изображения соответствующих ошибок печати, над которыми были проведены различные манипуляции, некоторыми из которых являются поворот, кадрирование, перевод изображения в цветовой режим «градации серого». Данные манипуляции над фотографиями, представленные на рис. 1, позволяют расширить датасет, снизить фактор влияния цвета детали при обучении нейронной сети и повысить качество машинного обучения.

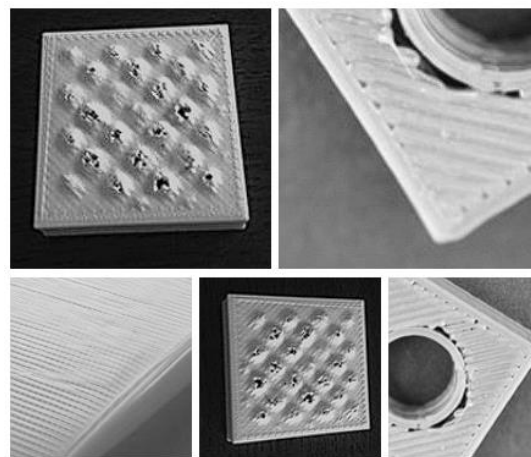


Рис. 1. Пример изображений категории
«Messy last Layer»

Для повышения точности распознавания аварийных ситуаций относительно результатов обучения первой версии нейронной сети только на наборе изображений был собран .csv файл. Он содержит вручную указанные координаты границ области ошибки, а также тип этой ошибки для каждого изображения из датасета. Одним из преимуществ данного подхода является выделение нескольких областей на одном и том же изображении, что расширяет размеченный набор данных. С помощью данного файла нейронная сеть будет искать на входном изображении участки с некорректной печатью.

Машинное обучение

Для нахождения участка с некорректной печатью на картинке был использован подход Faster RCNN с архитектурой ResNet50. Его архитектура состоит из нескольких шагов:

1. Определение набора областей, в которых предположительно находятся участки;
2. Извлечение из предполагаемых областей признаков с помощью свёрточной нейронной сети и их кодирование в вектор;
3. Классификация участка внутри области на основе вектора из пункта 2;
4. Улучшение координат предполагаемых областей (предполагаемая область обводится в цветную рамку). Для улучшения координат данной области размер рамки постепенно увеличивается, затем в

увеличенной области проверяется наличие участка с некорректной печатью;

5. Алгоритм повторяется, начиная со 2 пункта, пока не будут обработаны все области из пункта 1.

На выходе, после обучения мы получим модель в виде файла, в формате .pth, которая в последствии подключается к программному обеспечению.

Разработка программного обеспечения

Задачей программного обеспечения является анализ изображений с камер с помощью нейронной сети и обработка данных с датчиков. Алгоритм работы программы представлен на рис. 2.

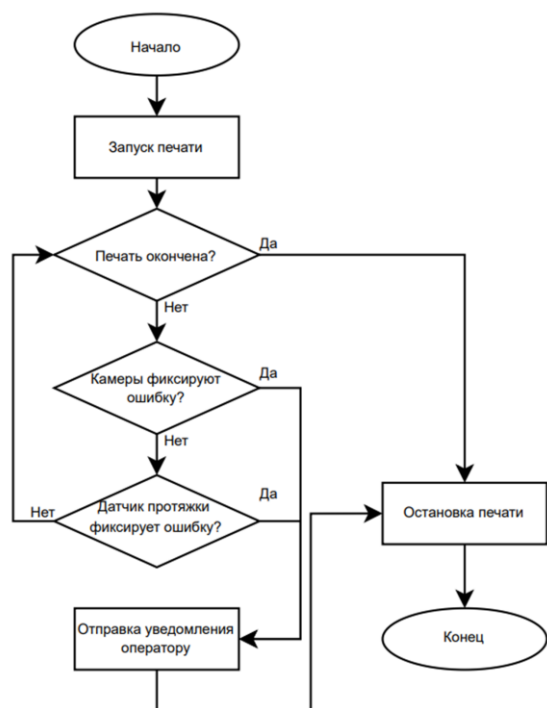


Рис. 2. Алгоритм работы программы

Получение данных осуществляется несколькими способами. Для определения некоторых возможных ошибок печати необходима информация с камер, которая, впоследствии, передается компьютеру для обработки [2]. Таким образом, для определения неправильной геометрии модели нейронная сеть проводит сравнение печатаемой детали с шаблонным изображением. Если изображения совпадают, то печатаемая деталь классифицируется как деталь без геометрических дефектов, в противном же случае оператор печати уведомляется об ошибке.

Для определения других ошибок может понадобиться более детальный анализ, например, для обнаружения обрыва нити подаваемого пластика. В этом случае работают датчики, которые посылают сигналы на плату Arduino Uno. Датчики могут фиксировать обрыв и окончание филамента, частичное расхождение модели, попадание сторонних объектов в область печати, а также такие дефекты как визуальные волны. Датчики протяжки на основе оптрона позволяют узнать об окончании или обрыве нити филамента: если оба датчика фиксируют прохождение нити, то подача филамента исправна. В против-

ных же случаях возникает обрыв нити или отсутствие материала.

Машинное обучение реализовано с помощью библиотеки ML.NET [3] на языке высокого уровня C# [4]. Данная часть программы является классификатором изображений, который обучен на ранее созданном датасете. Входными данными для нейронной сети являются изображения с камер, предварительно переведенные в тип byte. Результатом анализа нейронной сети является соотношение изображения с камер с одной из аварийных ситуаций или с отсутствием ошибок при печати.

Программа имеет два сценария работы: автоматическая передача изображений с камер для обработки нейронной сетью и загрузка конкретного изображения из файлов пользователя. В первом случае не требуется вмешательство пользователя программы для получения результата анализа, данные с камер будут подаваться с интервалом в 1 секунду, при этом информация об оценке соответствия процесса печати с одной из аварийных ситуаций будет своевременно выводиться на экран. Данный режим работы подойдет для оперативной проверки процесса печати на 3D-принтере. Во втором случае пользователь должен самостоятельно загрузить изображение, имеющееся на его компьютере. Этот режим подойдет для проверки корректности работы программы или при приостановленной печати.

Описанное программное обеспечение было разработано для 3D-принтера собственной разработки УЛ САПР МИЭМ НИУ ВШЭ. Возможно использование на других моделях 3D-принтеров, если их конструкция позволяет закрепить камеры и датчики для получения данных.

Оценка точности

Микроточность (MicroAccuracy) – это доля прогнозов, которые точно соответствуют истинным меткам класса. Приставка микро означает использование подсчета итоговых истинных положительных результатов, ложноотрицательных и ложноположительных результатов [5]. В диапазоне [0, 1], где 1 является максимальной точностью, микроточность обучения первой версии нейронной сети на датасете из изображений аварийных ситуаций составила 0,9304.

Заключение

Таким образом, с использованием дополнительного программного обеспечения становится возможным вовремя устранить аварийные ситуации, продлив при этом срок службы оборудования и сэкономив материалы.

Процесс устранения аварийной ситуации происходит с помощью уведомления, приходящего оператору печати, с просьбой остановить печать. Данное уведомление приходит в нескольких случаях:

- нарушение геометрии;
- обрыв нити;
- окончание филамента (печатающего материала);
- обнаружение различных ошибок при печати.

Использование методов машинного обучения позволяет обнаружить и классифицировать ошибку без непосредственного участия человека. Применение компьютерного зрения, внедрение и установка различных датчиков для считывания параметров печати способно зафиксировать аварийные ситуации, которые не охватывает датасет.

Список литературы

1. Техно 3D : Проблемы качества 3D-печати : [Электронный ресурс]. - 2013. - URL: <https://3dpt.ru/page/faq> (дата обращения: 26.01.2023).
2. Real Time 3D Printing Remote Defect Detection (Stringing) with Computer Vision and Artificial Intelligence, ResearchGate [Электронный ресурс]. URL: https://www.researchgate.net/publication/346995931_Real-Time_3D_Printing_Remote_Defect_Detection_Stringing_with_Computer_Vision_and_Artificial_Intelligence (дата обращения 26.01.2023)
3. Microsoft Learning : Документация по ML.NET: [Электронный ресурс]. - 2023. - URL: <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/> (дата обращения: 26.01.2023). - Текст : электронный.
4. Александрова И.Л. Программирование на языке C#. Учебное пособие / И.Л. Александрова, Д.Н. Тумаков. – Издание 2-ое, исправленное и дополненное. – Казань: Казанский государственный университет, 2017.
5. Microsoft Learning : Оценка результатов экспериментов с автоматизированным машинным обучением: [Электронный ресурс]. - 2023. - URL: <https://learn.microsoft.com/ru-ru/azure/machine-learning/how-to-understand-automated-ml?view=azureml-api-2> (дата обращения: 22.04.2023). - Текст : электронный.

РАЗРАБОТКА ПАРАМЕТРИЗИРОВАННОГО 3D ПРИНТЕРА

*Хачатрян А.А., Емельянова А.К.,
Михайловская А.Г., Нгуен А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Основная цель данной работы заключается в создании программы, которая может подстраивать все части, детали и характеристики 3D принтера под нужды заказчика. При изменении входных данных будет происходить параметризация ранее смоделированных в программе “Компас-3D” деталей и сборок.

Введение

В настоящее время 3D печать становится все более массовым явлением. Многие специалисты используют данную технологию для создания уни-

кальных изделий. Область 3D моделирования и печати актуальна во многих сферах жизни человека, таких как пищевое производство, строительство и медицина. Идея данной работы о параметризации 3D принтера заключается в том, что все характеристики принтера будут подстраиваться под нужды заказчика, что позволит значительно расширить вариативность изготавливаемых деталей, а также эффективно улучшить саму технологию печати.

Разрабатываемая программа подстраивает размеры всех деталей принтера под заданные параметры. Из измененных составляющих будет производиться сборка новой модели 3D принтера. Таким образом возможно изготовление объектов как маленьких размеров, так и крупных.

Аналоги разрабатываемой системы

В настоящее время не существует полностью параметризированной системы, способной изменять размеры принтера и его рабочей области, подстраиваясь под заданные пользователем параметры. На данный момент существует множество различных видов 3D принтеров, которые используются практически во всех сферах жизнедеятельности, однако не всегда имеется возможность найти подходящий 3D принтер ввиду того, что принтеры имеют определенные стандарты. Например, невыгодно приобретать принтер среднего размера, если ожидается, что нужно будет печатать только детали маленького размера. Так, заранее проделав настройку и указав необходимые данные, можно получить принтер оптимального размера с подходящими габаритами, сэкономив финансовые средства и материал, и получить 3D принтер, необходимый для конкретных целей. Такой подход является более рациональным, так как зачастую принтер нужен для определенных целей, и пользователь заранее знает, какими размерами и параметрами должен обладать принтер. Например, это применимо в производственных целях, когда необходимо печатать одинаковые модели в большом тираже, либо в учебных целях, когда заранее известны программа обучения и размеры печатаемых моделей, и следовательно габариты принтера.

Среда разработки

КОМПАС-3D [1] – мощная и универсальная отечественная система автоматизированного проектирования, позволяющая работать с проектами любой сложности, наглядно представить будущее изделие и проверить его собираемость, быстро подготовить документацию на изделие или объект, а также составлять чертежи и оформлять их в соответствии различным стандартам.

Ключевая особенность системы КОМПАС-3D – обеспечение сквозного проектирования от реализации идеи в 3D до подготовки полного комплекта документации. В связке с САМ-системами позволяет существенно сократить время подготовки изделия к производству. Имеет интеграцию с САЕ-системами, совместима с большинством САПР. Для быстрого освоения КОМПАС-3D обладает встроенными интерактивными учебными пособиями.

Система КОМПАС-3D включает в себя четыре основных компонента [2]:

- Чертежно-графический редактор (КОМПАС-График).
- Модуль проектирования спецификаций.
- Систему трехмерного моделирования.
- Текстовый редактор.

Чертежно-графический редактор КОМПАС-График позволяет работать с документами Чертеж, Фрагмент, Спецификация и Текстовый документ, а также получать данные из документов Деталь и Сборка.

Модуль проектирования спецификаций используется совместно с системой трехмерного моделирования и/или с редактором КОМПАС-График.

Система трехмерного моделирования КОМПАС-3D предназначена для создания трехмерных параметрических моделей деталей и сборочных единиц.

Текстовый редактор служит для создания различного рода текстовых документов.

Также имеется модуль Приложения, который содержит различные утилиты (Калькулятор, Комплектовщик документов КОМПАС-3D) и библиотеки (Материалы, Стандартные изделия, Сервисные инструменты и др.).

Все модули тесно интегрированы друг с другом. Справочники и прикладные библиотеки подключаются к системе по мере необходимости.

Разработка 3D моделей

Параметризация принтера подразумевает под собой то, что пользователь, указав необходимые габариты принтера и его рабочей области, получит на выходе готовую к печати 3D модель принтера нужного размера. Чтобы добиться такого результата, нужно сделать параметризацию большинства деталей, из которых собирается принтер, однако необходимо для начала разработать все модели составляющих принтера.

Разработка частей выполняется в программе КОМПАС-3D. Сборка принтера состоит из 39 уникальных деталей, большинство из которых подстраиваются под входные данные. Среди них: держатель вала, ролики, провода, рельсы, зубчатые шкивы. Все составляющие принтера представлены в виде 3D моделей, которые разработаны в соответствии с документацией и чертежами. Каждая из моделей, участвующих в параметризации, обладает переменными габаритами, которые меняются в зависимости от введенных данных. Например, длина рельсы напрямую зависит от площади рабочей области, а количество необходимых держателей вала в свою очередь связано с длиной самого вала.

Помимо параметризации размеров и количества деталей в зависимости от размеров принтера, также предполагается возможность изменения рабочего устройства 3D принтера. Еще одной функцией является наличие автоуровня для любого рабочего устройства. Автоуровень, или автокалибровка [3], это регулировка расположения столика для печати на протяжении всего процесса печати. Правильным расположением является такое расположение, когда

стол полностью параллелен плоскости осей X и Y, в которой собственно и перемещается печатающая головка принтера. При неправильном расположении возможны сильные искажения итоговой модели. Такую настройку можно производить и в ручную, например, подкрутив винты, однако, произведя автоматизацию данного процесса, можно сэкономить много времени и достигнуть большей точности.

На Рис. 1 представлен скриншот в программе КОМПАС-3D, на котором изображена одна из основных деталей итоговой сборки принтера, а именно экструдер E3D Titan AERO.

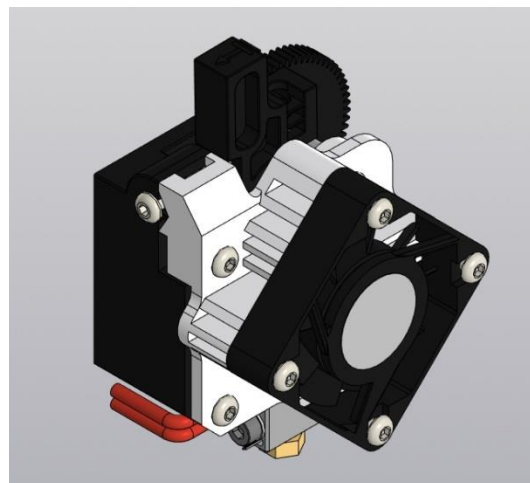


Рис. 1. Модель экструдера в программе Компас 3D

Программная часть

Одним из важных этапов работы является разработка программы, выполняющей роль моста между пользователем и готовой сборкой с возможностью изменения. Для выполнения этой задачи было решено использовать API Компас-3D. Архитектура программы и её связь с Компас-3D представлена на Рис. 2. Разработка включает в себя несколько этапов, начиная с параметризации каждого отдельного элемента (рис. 3), заканчивая целой сборкой, которая готова к работе с пользователем: на Рис. 4 и Рис. 5 представлена готовая сборка принтера после параметризации при различных входных данных.

После запуска программы и выбора необходимой сборки появляются текстовые поля, в которые надо вписать длину, ширину и высоту принтера. Если все данные входят в интервал допустимых значений (от 300 до 1000), то кнопка «Параметризовать» станет активной. Если какой-то из трех габаритов выйдет за пределы допустимых значений, кнопка «Параметризовать» будет неактивной и параметризация станет невозможной (Рис. 6). Для удобства детали были разделены по принципу их параметризации на группы:

1. Изменение положения в пространстве относительно начала координат (почти для всех);
2. Изменение длины, ширины, высоты детали (профили, винты, валы и т.д.);
3. Более сложные случаи (ремни, рельсы).

Сложные детали, как правило, состоят из многих объектов. Поэтому, чтобы не перегружать основную

сборку, было принято решение параметризовать такие объекты отдельно и затем добавлять в основную сборку. Примером является параметризация ремня, находящимся под столом принтера. Для его параметризации была написана функция «Pulley bottom». В ней используется 12 изменяемых переменных (6 точек, отвечающих за местоположение, и 6 радиусов, отвечающих за радиус). Каждая точка и радиус относятся к одному шкиву, на который натягивается ремень (Рис. 7).

Помимо этого, у каждой детали есть свой номер в сборке, через который происходит взаимодействие с ней.

Подробно изучив документацию к API [4], было решено использовать язык C# [5] из-за хорошей совместимости с кроссплатформенной Visual Studio, где помимо написания кода удобно реализована возможность создавать графический интерфейс, сделанный с применением технологии Windows Forms [6]. Необходимость разработки отдельной программы обусловлена тем, что ее функционал позволяет автоматически перестраивать всю сборку целиком с заданными параметрами. Если делать это в самом Kompas-3D, то надо будет высчитывать размеры и менять каждый компонент отдельно. Это довольно трудная задача, учитывая количество деталей в сборке. Также есть возможность отдельно сохранить сборку после параметризации для дальнейшей работы.

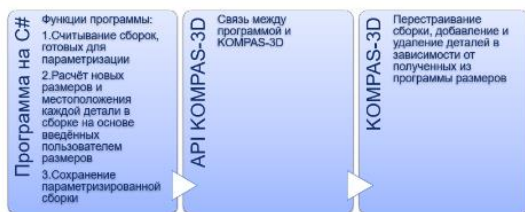


Рис. 2. Архитектура программы

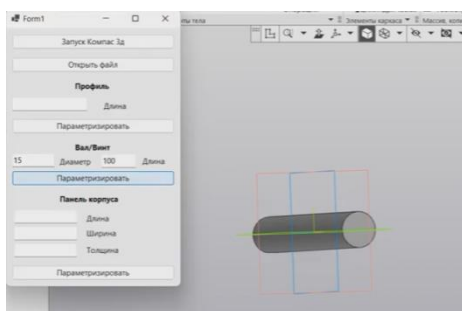


Рис. 3. Пример параметризации отдельной детали

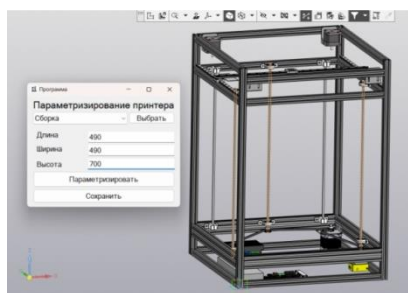


Рис. 4. 3D модель принтера после параметризации (1)

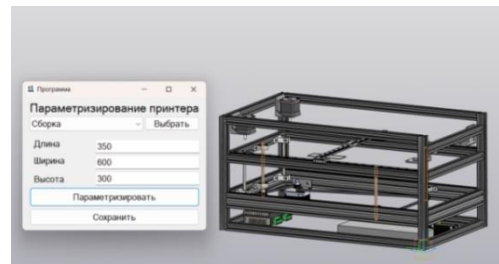


Рис. 5. 3D модель принтера после параметризации (2)

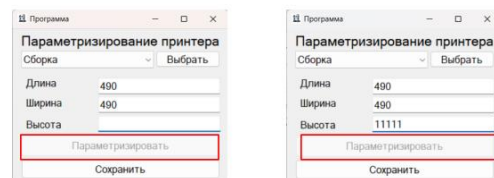


Рис. 6. Демонстрация неактивной кнопки «Параметризовать»

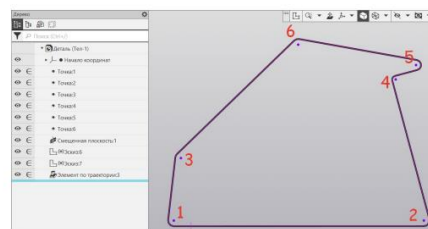


Рис. 7. Ремень, находящийся под рабочей областью принтера

Заключение

Учитывая растущую популярность 3D принтеров и 3D печати в целом, появляется необходимость в принтере, габариты которого заказчик сможет настроить самостоятельно. Разработанная программа позволяет подстраивать трехмерную модель, изменяя размеры своих составляющих в зависимости от входных данных, предоставленных пользователем.

Список литературы

1. Официальный сайт программы КОМПАС-3D, URL: <https://kompas.ru/>
2. Герасимов А. А. Самоучитель КОМПАС-3D V19. – СПб.: БХВ-Петербург, 2021. – 624 с.: ил. – (Самоучитель)
3. Разновидности систем автокалибровок стола 3D принтера, URL: <https://3d-diy.ru/wiki/3d-printery/raznovidnosti-sistem-avtokalibrovki-stola-3d-printera/>
4. Справочная система КОМПАС-МАСТЕР, URL: <https://kompas.ru/kompas-3d/application/machinery/spravochnik-konstruktora/>
5. Документация по C#. Начало работы, руководства, справочные материалы | Microsoft Learn [Электронный ресурс]. URL: <https://learn.microsoft.com/ru-ru/dotnet/csharp/>
6. Обзор - Windows Forms .NET Framework | Microsoft Docs [Электронный ресурс]. URL: <https://docs.microsoft.com/ru-ru/dotnet/desktop/winforms/windows-formsoverview?view=netframeworkdesktop-4.8>

3D-РЕКОНСТРУКЦИИ И МОДЕЛИРОВАНИЕ ИСТОРИЧЕСКОГО КОНТЕНТА ДОМА ДУРАСОВЫХ

*Цуркан А.А., Старилова Е.В., Нигматулин Н.Р.,
Сватухин А.М., Жибуртович Е.М.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе представлено описание разработки виртуального гида для Дома Дурасовых. В процессе исследования было разработано приложение, а также создан 3D контент в соответствии с различными историческими эпохами здания для виртуальных экскурсий.

Введение

Современный рынок технологий предоставляет решения как для крупных, так и для небольших музеев [1], которые могут преследовать три цели [2]:

- осуществление регистрации, хранения и учета музейных коллекций;
- повышение интереса посетителей к выставкам путем лучшей репрезентации коллекций;
- воссоздание исторического облика различных сооружений и экспонатов.

Подобные разработки применяются для создания глобальных платформ для доступа к историческим объектам из любой точки мира при наличии практически любого устройства с доступом в интернет.

В период пандемии COVID-19 посещение музеев стало невозможным в связи с введением ограничительных мер. В это время стал наблюдаться рост интереса к виртуальным выставкам, сделанным на основе 360-туров. Фактически они остались единственным способом посещения музеев в период распространения коронавирусной инфекции [3].

Даже если ограничения по COVID-19 уже сняли, то возможность побывать на самых знаковых выставках мира и увидеть легендарные произведения искусства остается. Если у вас нет уникальной возможности оказаться в Лувре и вы не можете в живую увидеть «Мона Лизу», то это не означает, что вы вообще не можете ее оценить. Именно виртуальные музеи могут помочь рассмотреть картину со всех сторон. Лувр также предлагает взглянуть на то, как выглядел в средневековье, когда изначально был крепостью, построенной королем Филиппом Огюстом. Получается, что виртуальные выставки дают возможность увидеть не только текущий вид музея, но и его прошлые облики. Многие известные музеи уже используют технологии для создания 360-экскурсий, что говорит о том, что они набирают большую популярность.

Целью работы является 3D-реконструкция и моделирование различных исторических периодов Дома Дурасова, а также создание виртуального гида с использованием смоделированного исторического контента.

Для достижения цели были поставлены следующие задачи:

- Создание интерактивной трехмерной структуры залов Дома Дурасова вместе с расположенными в нем экспонатами.
- 3D-моделирование пространств и объектов различных исторических периодов Дома Дурасова, недоступных для оцифровки методами фотограмметрии или лазерного сканирования.
- Наполнение проекта текстовыми описаниями, аудиосопровождением, создание интерактивных анимированных экскурсий.
- Локализация виртуального тура на несколько языков для привлечения иностранной аудитории.
- Интеграция разработанных 3D-моделей и программных модулей с веб-выיוером и единым веб-порталом Дома Дурасова.

3D контент

Одним из основных направлений данного проекта является создание базы 3D моделей для дальнейшего их использования как в веб-выювере, так и в смежных проектах.

Трёхмерные сцены разрабатывались в двух направлениях. Первое направление разработки сцен - визуализация сцен без масштабирования. Второе направление – разработка сцен с возможностью масштабирования (модель здания Дома Дурасовых).

В первом направлении сцен разрабатывались предметы мебели и интерьера в рамках эпохи практической академии экономических наук. Модели создавались по архивным фотографиям, без привязки к размерам и с возможностью масштабирования во вью-порте программы для сборки полноценных сцен. Каждый участник проекта работал в удобном для себя ПО, из-за чего цикл разработки моделей (далее пайплайн) менялся в зависимости от программы, однако до этапа создания текстурных карт и разверток он был один:

1. Драфт - создание упрощенной формы и силуэта модели из примитивных элементов программы (куб, цилиндр, конус и т. д.)
2. Создание LowPoly модели – легкая форма финального варианта, в которой оптимизируется геометрия, невидимые полигоны удаляются.
3. Создание HighPoly модели – самая детализированная форма модели, которая не подразумевает ограничений в количестве полигонов и точек модели.

Далее, модель, если до этого момента она разрабатывалась в системе автоматизированного проектирования (САПР), передавалась участнику, работающему непосредственно в программах с полигональным типом моделирования, где происходила подготовка к запеканию текстур, создание развертки и текстурирование.

Помимо САПР и полигонального моделирования, также применялся метод скульптинга для придания необходимого реализма объектам. Скульптинг – способ моделирования, который применяется для создания преимущественно органических форм и моделей с повышенной детализацией, либо для детализации

имеющейся модели. С помощью скульптинга можно воссоздать складки на ткани, текстуру древесины и подобные формы. В проекте для скульптурирования были использованы программы ZBrush и Blender.

Полигональное моделирование – способ моделирования, который чаще всего применяется для создание искусственных твердотельных объектов. В проекте были использованы Autodesk 3DsMax, Blender (включает в себя возможности как полигонального моделирования, так и скульптинга).

Кроме того, производился подбор материалов, которые могли существовать и использоваться в той или иной эпохе. Модели, созданные в рамках проекта представлены на рисунках 1-3. Создание текстурных координат происходило в самой программе, каждый использовал индивидуальный способы, так как это не влияло на качество работы, однако значительно сокращало затраченное время 3D модельерами на освоение новых программных обеспечений.



Рис. 1. 3D модели мебели исторического периода практической академии экономических наук

Для тестирования качества текстурирования моделей и вида 3D сцен была использована профессиональная платформа для рендеринга KeyShot 5 совместно с ведущей в отрасли программой набора инструментов для рендеринга, анимации и запекания текстур в реальном времени — Marmoset.

Для рендера панорам в 3DsMax использовался специальный модификатор CoronaCamMod, который позволял выставить угол обзора и рендера вплоть до 360 градусов.

В рамках второго направления сцен (с возможностью масштабирования) была переработана модель здания Дома Дурасовых, добавлены недостающие проходы и окна. На данный момент реализуется переработка под каждую эпоху в отдельности, так как здание имело свои особенности в каждый момент времени.

Использование готовой модели планируется не только в виртуальном гиде и презентации дома с анимацией показа внутреннего устройства усадьбы в каждый период времени, но и для создания физического макета Дома Дурасова. На данный момент завершается этап печати и первичной обработки изде-

лия. Печать велась на 3D-принтерах с использованием технологий FFF, FDM и SLA. Обе технологии, упомянутые первыми, схожи по реализации. Экструдер плавит нить материала (филамент) и укладывает его слоями, образуя заданную программой модель. В свою очередь принтеры, оборудованные технологией SLA печати, используют в качестве филамента жидкие фотополимерные смолы. Отверждение происходит за счет отображения на матрице с ультрафиолетовыми светодиодами маски (черно-белого изображения), обозначающей, какая форма слоя застынет. SLA-принтеры отличаются не только способом наращивания слоев, но и минимальной высотой слоя, которая может достигать до 0.025-0.05 мм. Данный параметр отличается от принтера к принтеру и напрямую зависит от разрешения матрицы в основании принтера.



Рис. 2. 3D модель пространства крепостного театра в усадьбе Дурасова



Рис. 3. 3D мебель для интерьера периода практической академии экономических наук

Специфика приложения виртуального тура по дому Дурасовых

Одним из главных аспектов в реконструкции исторического контента дома Дурасовых является создание виртуального гида.

Следует отметить, что при разработке виртуального тура была учтена специфика Дома Дурасовых:

расположение точек интереса внутри здания, необходимость показать все 4 исторических периода существования дома Дурасовых, наличие встроенного аудиогuida и текстовых описаний для точек интереса, а также необходимость локализации тура на несколько языков для привлечения иностранной аудитории.

С учетом вышеупомянутой особенности Дома Дурасовых, был проведен обзор и анализ существующих наборов инструментов для создания виртуальных туров на российском и мировых рынках. Результатом стало решение выбрать платформу для создания виртуальных туров krapo, которое бы удовлетворяло всем ожидаемым требованиям: наличие аудиогuida и текстового описания для каждой точки интереса [4].

Описание разработанной информационной системы

С учетом анализа существующих аналогов и технического задания, был выделен необходимый функционал системы:

- единый каталог виртуального музея с рубрикаторм по пространственному или тематическому принципу;
- возможность переключаться между различными историческими периодами дома Дурасовых;
- настройки приложения для выбора языка;
- аудиоплеер для аудиогuida.

Для разработки виртуального тура использовался язык разметки XML. Для интеграции тура используется язык JavaScript. При разработке единого каталога использовался API WebGL для просмотра 3D объектов в браузере [5].

Тур состоит из 6 модулей: tour.html, tour_start.xml, tour_ru.xml, tour_en.xml, tour_fr.xml и vtourstkin.xml. Модуль tour.html используется для запуска и встраивания тура. Модуль tour_start.xml отвечает за стартовое окно выбора языка, так же он описывает методы для смены языка внутри тура. Модули tour_ru.xml, tour_en.xml, tour_fr.xml описывают основные функции приложения и отвечают за русский, английский и французский языки соответственно. Модуль vtourstkin.xml описывает объекты интерфейса и методы для переключения между сценами.

Описание работы приложения

При входе пользователю необходимо выбрать язык (рис. 4).



Рис. 4. Окно выбора языка

После выбора языка пользователь попадает на начальную точку тура (рис. 5).

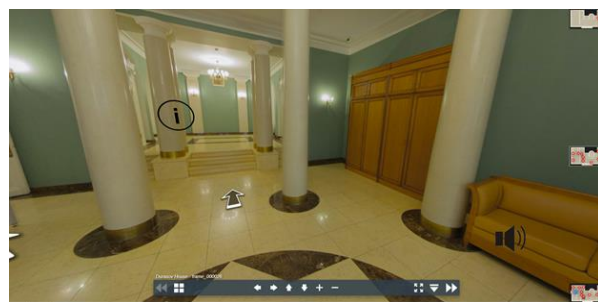


Рис. 5. Начальная точка тура

Информационная метка открывает для пользователя информацию о точке интереса (рис. 6). Чтобы закрыть информационное поле нужно нажать на него.

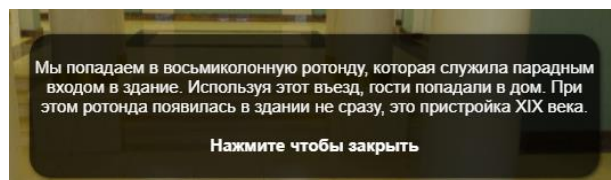


Рис. 6. Окно с информацией о точке интереса

Справа на экране присутствует план текущего этажа здания (рис. 7). На планах установлены метки. При нажатии на них пользователь перейдет к соответствующей точке интереса. Также можно перейти к точкам интереса, находящимся на других этажах, переключив план при помощи кнопок над картой. Выбранный этаж подсвечивается белым цветом.



Рис. 7. Планы этажей

На нижней панели представлены (рис.8) 11 кнопок. Кнопки 1-2 переключают панорамы в заданном порядке. Кнопки 3-6 дублируют управление камерой. Кнопки 7-8 отвечают за приближение камеры. Кнопка 9 открывает меню со всеми точками интереса тура. Кнопка 10 разворачивает тур на полный экран. Кнопка 11 сворачивает панель управления.



Рис. 8. Панель управления

Заключение

Таким образом, был разработан виртуальный интерактивный тур - приложение, рассказывающее историю, с помощью моделирования различных исторических периодов Дома Дурасовых.

В рамках данной работы были разработаны модели внутреннего убранства Дома Дурасовых, включающие в себя мебель, лепнины, предметы интерьера и др. Были воссозданы комнаты со смоделированными объектами, повторяющие сохранившиеся архивные фотографии реальных комнат прошлых эпох Дома Дурасовых. Для моделирования были использованы такие программы, как 3Ds Max, Blender, ZBrush и Fusion360. Приложение разрабатывалось на языке xml с использованием средств программы KRpano.

Список литературы

1. Halyna Lypak, Nataliia Kunanets, Volodymyr Pasichnyk, Nataliia Veretennikova, Digitization Project for Historical and Cultural Heritage // 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT) - T.2. - 2020. - С.194-198.
2. Ramy Hammady, Minhua Ma, Nicholas Temple Augmented Reality and Gamification in Heritage Museums // Serious Games: Second Joint International Conference, JCSG 2016. – Т. 9894. - 2016. - С. 181–187.
3. Giannini T., Bowen J. P. Museums and Digital Culture: From Reality to Digitality in the Age of COVID-19 // Heritage. - 2022. - Т. 5. - № 1. - С. 192–214.
4. Krpano XML reference [сайт]. - URL: [https://krpano.com/docu/xml/\(дата обращения: 25.01.2022\)](https://krpano.com/docu/xml/(дата обращения: 25.01.2022)). - Текст: электронный.
5. WebGL // 2D and 3D graphics for the web [сайт]. URL: mzl.la/3iJA4e0 (дата обращения: 26.01.2022). - Текст: электронный.

РАЗРАБОТКА САПР ДЛЯ ВЫСОКОУРОВНЕВОГО МОДЕЛИРОВАНИЯ СЕТЕЙ НА КРИСТАЛЛЕ

*Бородин Н.Ю., Пономарев А.С., Таржанов Т.В.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе рассматриваются ключевые моменты разрабатываемой САПР для моделирования сетей на кристалле (далее – СтнК). Применяемые архитектурные решения позволяют упростить процесс разработки и поддержки САПР, улучшить ее пригодность для использования на различных платформах. Методы оптимизации повышают эффективность моделирования путем его ускорения, уменьшения вероятности возникновения критических ошибок и увеличения точности расчетов характеристик. Модуль поддерж-

ки принятия решений поможет повысить эффективность выбора пользователем моделей и методов для расчета заданных характеристик СтнК.

Введение

Сложность решаемых информационными системами задач, а также объем обрабатываемых данных с каждым годом увеличиваются, соответственно растут и требования к производительности информационных систем с целью уменьшения времени обработки информации. Улучшение производительности информационных систем методом увеличения плотности транзисторов уже стало невозможным, и это открыло путь к эволюции информационных систем на основе многоядерности и многопоточности [1].

Важным компонентом систем на кристалле (далее – СтнК), существенно влияющим на эффективность работы, является СтнК – подсистема связи между ядрами и различными узлами данных на интегральной схеме. В связи с этим возникает ответственная задача по проектированию СтнК, для решения которой необходимы специализированные инструменты [2].

Процесс разработки СтнК

Разработка СтнК проходит в несколько взаимосвязанных последовательных стадий, как показано на рис. 1. На каждом из изображенных этапов применяются различные программные средства.

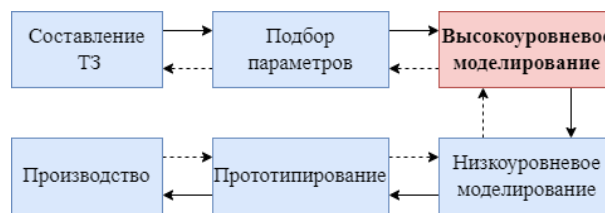


Рис. 1. Этапы разработки СтнК

В данной работе рассматривается стадия высокоуровневого моделирования СтнК и разработка программного обеспечения для его проведения.

Предпосылки создания САПР для высокоуровневого моделирования СтнК

В настоящее время существует большое количество программ, называемых высокоуровневыми моделями СтнК, которые позволяют по заданным параметрам рассчитать ожидаемые характеристики. Но использование отдельно взятых моделей имеет ряд недостатков. Так, каждая модель способна обрабатывать только немногие входные параметры и выходные характеристики. Кроме того, на определенных наборах входных данных модель может возвращать некорректные результаты. Также модели часто не предоставляют пользователям встроенных средств для оптимизации процесса моделирования, анализа результатов и сравнения результатов. Стоит отметить, что большинство моделей не имеет веб-версий и требует установки дополнительного программного обеспечения на компьютер пользователя.

Перечисленные выше проблемы привели к появлению идеи реализации САПР для высокоуровневого моделирования СтнК без упомянутых недостатков.

Концепция САПР для высокоуровневого моделирования СтнК

Предлагается создать САПР, состоящую из клиентского приложения, серверного приложения, включающего в себя несколько логических компонентов, и базы данных.

Клиентское приложение должно представлять собой легковесный исполняемый файл, не требующий установки дополнительного программного обеспечения. Данное приложение должно позволять настраивать желаемые расчеты, вводить параметры СтнК и отображать полученные характеристики.

Непосредственно расчеты и взаимодействие с моделями должно осуществлять серверное приложение, с которым клиентское приложение взаимодействует по API. Важными компонентами серверного приложения являются блок оптимизации процесса моделирования и система поддержки принятия решений, принцип работы которых подробнее представлен в соответствующих частях данной работы. Еще одним существенным элементом серверного приложения является блок взаимодействия с моделями, который должен быть реализован таким образом, чтобы подключение к САПР различных моделей не требовало внесения изменений в код системы.

Использование базы данных, взаимодействующей с серверным приложением, позволит накапливать результаты проводимых моделирований, что даст возможность уменьшить количество проводимых вычислений и впоследствии даст возможность проводить анализ данных.

Методы оптимизации моделирования

Для оценки эффективности моделируемой СтнК на этапе высокоуровневого моделирования необходимо найти такую частоту генерации данных, при которой достигается максимально возможная пропускная способность сети. Таким образом возникает необходимость оптимизировать процесс нахождения этой частоты с целью повышения точности вычисления и уменьшения затрачиваемого времени [3].

Одним из реализованных методов оптимизации в САПР является использование 3 различных алгоритмов: пошаговый поиск, метод дихотомии и метод золотого сечения.

Пошаговый поиск [4]. К входным данным алгоритма относятся начальная частота генерации данных, количество итераций и «шаг» - величина, на которую будет увеличиваться частота каждую итерацию. Данный алгоритм достаточно точен и прост в реализации, но при моделировании СтнК со сложной топологией и высоким числом узлов сети общее время моделирования с использованием будет существенно больше, чем при применении других алгоритмов. Принцип работы пошагового поиска представлен на рис. 2.

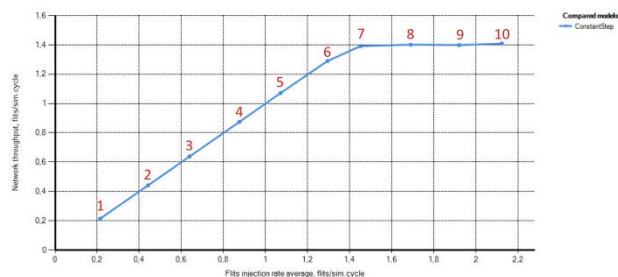


Рис. 2. Пошаговый поиск

Метод дихотомии [5]. К входным данным алгоритма относятся минимальная и максимальная частоты генерации данных и количество итераций. На каждой итерации путем вычислений находится середина между минимальной и максимальной частотами, после чего определяется, находится полученная точка до или после искомой частоты. На основании этой информации изменяется либо минимальная, либо максимальная из частот на вычисленное значение. Данный подход позволяет достичь малых значений погрешности нахождение искомого значения за небольшое количество итераций даже в случае сложной топологии сети и большого количества узлов. Принцип работы метода дихотомии представлен на рис. 3.

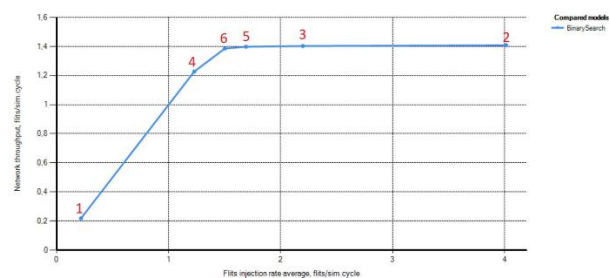


Рис. 3. Метод дихотомии

Метод золотого сечения [6]. К входным данным алгоритма относятся минимальная и максимальная частоты генерации данных и количество итераций. Принцип работы данного алгоритма похож на метод дихотомии с тем отличием, как вычисляется следующая частота. В методе золотого сечения вычисляется не середина, а 2 точки по формулам золотого сечения, и затем определяется, в каком из 3 получившихся отрезков находится искомая частота. Принцип работы метода золотого сечения представлен на рис. 4.

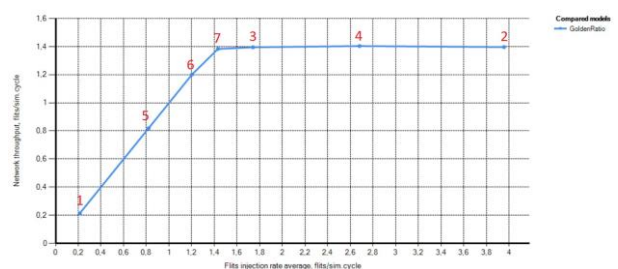


Рис. 4. Метод золотого сечения

На основе результатов сравнения числа необходимых итераций для каждого из алгоритмов был сделан вывод, что в одной из заданных топологий сети алгоритмы поиска методом дихотомии и методом золотого сечения являются наиболее эффективными с точки зрения минимизации количества итераций и затрачиваемого времени для достижения одной и той же погрешности определения точки насыщения.

Система поддержки принятия решений

Взаимодействие с САПР рассчитано на пользователей с различным уровнем знаний в области СтнК, ими могут быть как разработчики, так и те, кому нужно просто получить наилучшую модель и ее результат для заданных входных параметров, при этом не обладая специальными навыками. Для последнего случая и предполагается создание модуля системы поддержки принятия решений (далее - СППР). СППР – это компьютерная автоматизированная система, основной задачей которой является помощь людям, принимающим решения в сложных условиях для полного и объективного анализа предметной деятельности [7]. У такой помощи существует два основных направления:

- облегчение взаимодействия между данными, процедурами анализа и обработки данных и моделями принятия решений, с одной стороны, и с лицом, принимающим решения – с другой;

- предоставление вспомогательной информации, в основном для решения слабоструктурированных или неструктурированных задач, для которых трудно заранее определить данные и процедуры соответствующих решений.

В САПР разрабатываемая СППР поможет повысить эффективность выбора пользователем моделей и методов для расчета заданных характеристик СтнК. Для ее создания предстоит определить зависимость входных и выходных параметров моделей высокоуровневого моделирования СтнК. Сценарием использования разрабатываемой СППР является ввод входных параметров моделей СтнК с указанием важных для расчетов критериев, в результате будет получено множество наилучших моделей по указанным критериям.

Заключение

Применение САПР для высокоуровневого моделирования СтнК со встроенными методами оптимизации ускоряет процесс моделирования и повышает степень его автоматизации. СППР позволяет понизить порог вхождения пользователей для работы с моделями СтнК и повысит эффективность выбора моделей и методов для расчета заданных характеристик СтнК. Применяемые архитектурные решения упростят разработку и поддержку системы, а также облегчат процесс добавления новых высокоуровневых моделей в САПР.

Список литературы

1. Недбайло, Ю.А. Разработка сети на кристалле для перспективных многоядерных микропроцессоров

/ Ю.А. Недбайло // Труды Московского физико-технического института. – 2017. – Т. 9. – № 2 (34). – С. 151–163.

2. Американов А. А. Автоматизация высокоуровневого моделирования сетей на кристалле // В кн.: Проблемы разработки перспективных микро- и наноэлектронных систем – 2021 (МЭС-2021) / Под общ. ред.: А. Л. Стемпковский. Вып. 1. ИППМ РАН, 2021. С. 39-45.

3. A.A. Amerikanov, A.S. Ponomarev, T.V. Tarzhanov. Automation of NoC throughput search in high-level modeling, in: Proceedings of 2022 IEEE Moscow Workshop on Electronic and Networking Technologies (MWENT) / Сост.: И. А. Иванов, О. Стукач.; Ed. by O. Stukach. M.: IEEE, 2022. P. 1-5.

4. Самуйлов С. В. Методика сравнительного анализа алгоритмов на примере алгоритмов последовательного поиска // Концепт. – 2014. – №. 9. – С. 46-50.

5. Ткачева Е. А. Алгоритм бинарного поиска // Электронное обучение в непрерывном образовании. – 2014. – №. 1-2. – С. 341-343.

6. Альхимович М. А. Метод золотого сечения для решения задач минимизации. – 2017.

7. Медведева М. А., Коломыцева А. О., Вишнякова А. Ю., Искра Е. А. Системы поддержки принятия управленческих решений - Екатеринбург: УрФУ, 2019.

ПРОЕКТИРОВАНИЕ СРЕДЫ МОДЕЛИРОВАНИЯ КАНАЛОВ И СЕТЕЙ СВЯЗИ ПОДКЛЮЧЕННОГО ТРАНСПОРТА

*Степанянц В.Г., Андрончев А.Д.,
Шабалин Е.А., Добжанский Е.К.*

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В последние несколько лет активно исследуются алгоритмы совместного восприятия для подключенного беспилотного транспорта. Программное обеспечение, требующееся для их тестирования с помощью компьютерного моделирования, находится на начальных этапах разработки и не включает в себя модели распространения радиосигналов. В работе проводится анализ особенностей моделирования в области совместного восприятия. В результате анализа определены требования к модулю интегрированной среды моделирования, направленному на выполнение поставленной задачи. Проведен обзор существующих подходов и инструментов. На основе полученных требований спроектирована архитектура модуля среды моделирования каналов и сетей связи подключенного транспорта.

Введение

Более 1,35 миллиона человек в год погибают в дорожно-транспортных происшествиях, и более 50 получают травмы или становятся инвалидами на всю жизнь [1]. Решения для этой проблемы разрабатываются в области интеллектуальных транспортных систем (ИТС), беспилотного транспорта, подключенного транспорта, и других технологий управления дорожным движением. Беспилотные автомобили планируют траекторию движения, основываясь только на информации, которую они получают от установленных на них сенсоров. Их возможности по планированию траектории ограничены дальностью восприятия датчиков. Подключенные автомобили обмениваются информацией с помощью всенаправленных радиосигналов и могут обмениваться сообщениями, содержащими их местоположение и другую информацию без очевидных слепых зон в пределах их диапазона связи. Технологии беспилотного и подключенного транспорта могут дополнять друг друга. В таком случае реализуется концепция CDA (Cooperative Driving Automation): беспилотные автомобили могут обмениваться информацией между собой, с дорожной инфраструктурой и другими участниками дорожного движения с помощью беспроводной связи. Поскольку для подключенного беспилотного транспортного средства таким образом доступно больше информации о дорожной обстановке, оно может использовать ее для уточнения планирования движения. CDA привлекает все больше внимания в последние несколько лет [2].

Перед тем, как такие системы можно будет использовать, они должны пройти обширное тестирование. Возможности по тестированию CDA-технологий в реальном мире и на дорогах общего пользования крайне ограничены. Вместо этого, для изучения их влияния на транспортную систему может использоваться высокоточное компьютерное моделирование. Разработка сред моделирования для алгоритмов совместного восприятия дорожной ситуации подключенным беспилотным транспортом находится на начальных этапах, но уже существуют интегрированные среды моделирования высокой точности. Например, среда OpenCDA [3], соединяющая микроскопический симулятор транспортных потоков SUMO и графический симулятор окружающей среды для беспилотного транспорта CARLA. Для моделирования некоторых алгоритмов совместного восприятия необходимо дополнительно учитывать задержку при передаче информации между участниками дорожного движения [4]. Такая возможность реализована в OpenCDA только в виде выбора предустановленного или случайного времени задержки, что не является достаточно точным описанием рассматриваемой системы и может влиять на результаты моделирования.

Целью работы является проектирование архитектуры модуля среды моделирования алгоритмов совместного восприятия, позволяющего учитывать задержку передачи сигналов, распространяющихся по беспроводным каналам самоорганизующейся сети

связи, соединяющей участников движения (Vehicular Ad-hoc NETwork - VANET-сеть).

В дальнейших разделах работы, соответственно, будут выполнены следующие задачи:

- определение особенностей моделирования рассматриваемых технологий и соответствующие им требования к программной реализации среды моделирования каналов и сетей связи подключенного транспорта;
- обзор и анализ существующих подходов к моделированию и использующихся инструментов;
- проектирование архитектуры модуля среды моделирования алгоритмов совместного восприятия, позволяющего учитывать задержку передачи сигналов подключенного транспорта.

Особенности моделирования каналов и сетей связи подключенного транспорта

Современные симуляторы VANET прошли долгий путь от ранних сред моделирования VANET, которые часто использовали упрощенные нереалистичные модели, такие как случайное перемещение транспортных средств, среда без помех или дальность передачи, ограниченная заданным радиусом. Тем не менее, еще предстоит приложить значительные усилия для повышения реалистичности и эффективности симуляторов VANET. В этой работе мы выделяем три ключевых составляющих VANET-симуляторов [5]:

- Модели мобильности описывают реалистичное движение транспорта, включая ограничение движения транспортных средств проезжей частью, взаимодействие между транспортными средствами;
- Сетевые модели разработаны для описания реалистичного обмена данными, включая имитацию контроля доступа к среде (MAC), маршрутизации и протоколов верхнего уровня;
- Проектирование подключенного транспорта сопряжено с определенными трудностями, и наиболее сложной задачей является учет влияния внешней среды и ее объектов на распространение сигнала по беспроводным каналам связи [6]. Модели распространения радиосигнала направлены на реалистичное моделирование сложной среды, включая как статические объекты (например, здания, эстакады, деревья, профиль местности), так и мобильные объекты (другие транспортные средства на дороге), окружающей подключенные транспортные средства. Полевые эксперименты показывают, что наличие препятствий для сигнала и движение приемопередатчиков приводят к потере передаваемой информации [7].

Модели мобильности уже включены в среду моделирования OpenCDA. Для создания моделей VANET-сетей существует множество симуляторов разной точности. При заданной модели мобильности точность обуславливается моделью распространения радиосигналов. Упрощенные стохастические модели радиосвязи (например, модели свободного пространства, логарифмических потерь на расстоянии, двухлучевого отражения от земли и т.д.), которые основаны на статистических свойствах выбранной среды

и не учитывают конкретные препятствия в рассматриваемом регионе, не могут обеспечить удовлетворительную точность для типичных сценариев VANET. В отличие от них, реалистичные геометрические модели каналов (например, основанные на трассировке лучей), учитывающие топографию местности, дают результаты, которые хорошо согласуются с реальным миром [5]. Таким образом, модуль для учета задержки при передаче сигналов в VANET-сетях должен включать в себя сетевую модель и геометрическую модель распространения сигналов.

Обзор существующих подходов и инструментов для моделирования

Существует ряд сетевых симуляторов, например, ns-2, ns3, OMNeT++, SWAN, OPNET, Jist, GloMoSim и т.д. В настоящее время, особенно в автомобильных приложениях, наиболее часто используются два симулятора – OMNeT++ и ns3. Оба этих симулятора имеют открытый исходный код. При сравнении ns3 значительно сильнее нагружает процессор при моделировании того же количества узлов (примерно 10-кратная разница при моделировании сотен узлов) [8]. Это важно при использовании сетевого симулятора как части интегрированной среды моделирования, в которой несколько симуляторов работают одновременно.

В настоящее время разработчиками поддерживаются следующие симуляторы и интегрированные среды моделирования для моделирования VANET с открытым исходным кодом: Veins, Eclipse MOCAIC (ранее коммерческий симулятор VSimRTI) [9].

Eclipse MOCAIC – это фреймворк, который позволяет взаимодействовать с различными симуляторами. Он поддерживает симуляторы мобильности: SUMO, VISSIM и PHABMACS, а также сетевые симуляторы: ns3, OMNeT++, SNS, Eclipse MOCAIC Cell. Также, возможна интеграция с другими симуляторами, такими как CARLA.

Veins основан на OMNeT++ и SUMO. Veins – это более зрелый фреймворк, и в настоящее время с его использованием было проведено больше исследований, для него существует больше расширений [9]. Например, фреймворк INET, который используется для моделирования сетей связи. Другим расширением Veins является Artery [10]. Он был разработан для поддержки европейских стандартов связи подключенного транспорта.

Геометрическая модель в свою очередь может быть стохастической или детерминированной [11]. Стохастические геометрические модели используют упрощенное представление распространения сигналов в среде. Детерминированные геометрические модели распространения характеризуют канал связи, используя подробную информацию о среде. Более вычислительно легкой альтернативой решению уравнений Максвелла для радиоволн, распространяющихся в среде, является использование детерминированных геометрических методов трассировки лучей, геометрической оптики и геометрической теории дифракции. Точность этого подхода зависит от

строгости описания среды распространения и всех объектов, которые могут влиять на распространение радиоволн.

Мы выделяем квазидетерминированную геометрическую модель распространения GEMV2 [12], которая использует границы транспортных средств, зданий и листья, чтобы различать три типа связи: прямая видимость, отсутствие прямой видимости из-за транспортных средств и из-за статических объектов. Для каждого канала GEMV2 вычисляет крупномасштабное затухание сигнала детерминировано, а мелкомасштабное затухание стохастически на основе количества и размеров окружающих объектов. Такой подход сохраняет точность и при этом позволяет сэкономить вычислительную мощность. GEMV2 реализована для фреймворка INET и имеет открытый исходный код.

Архитектура модуля связи среды моделирования алгоритмов совместного восприятия

Далее, на основе определенных требований и рассмотренных инструментов, мы проектируем архитектуру модуля для учета задержек при передаче информации для моделирования совместного восприятия. При разработке программы необходимо обеспечить возможность использования следующих видов симуляции: фреймворк Artery, разработанный на основе OMNeT++ и модель каналов связи GEMV2. Данные инструменты выбраны на основании проведенного анализа и того, что Artery использует трафик, создаваемый в SUMO, и имеет в себе реализацию модели GEMV2. Таким образом, единственными доработками для разработки среды моделирования являются создание модуля «CARLA-OMNeT++ bridge» для передачи данных, полученных сенсорами в симуляторе CARLA в модули сенсоров и сервисов Artery, и доработка этих модулей для хранения и обработки новых типов данных.

На рис. 1. представлена архитектура спроектированного модуля. Прямоугольниками обозначены основные программные модули среды моделирования. Зеленым цветом обозначены модули, обрабатывающие информацию о дорожно-транспортной сети и потоке транспорта, голубым – обрабатывающие информацию об окружающих объектах в 3D-среде CARLA, бирюзовым – моделирующие обмен информацией по беспроводным каналам подключенного транспорта. Стрелками обозначены направления передачи данных между программными модулями. SUMO, являющийся частью OpenCDA, управляет трафиком (модуль Vehicle Traffic) и светофорами (Traffic Lights), создает дорожную сеть (Road Network). SUMO включает в себя интерфейс TraCI, который позволяет обмениваться информацией с другими симуляторами. CARLA предоставляет информацию о данных, получаемых с сенсоров. Эта информация передается через интерфейс в Artery, где распространяется по внутренним модулям. Модуль Radio, описывающий канал связи, использует модель GEMV2 через INET.

Таким образом, спроектированный модуль среды моделирования позволяет добиться по сравнению с

фреймворком OpenCDA следующих качественных свойств: учет в модели объема передаваемых данных, мощности передатчиков, затухания сигналов при распространении в свободном пространстве и через препятствия. Данные свойства позволяют получать реалистичную информацию о том, какие сообщения будут переданы между транспортными средствами, а какие нет (в OpenCDA передаются все сообщения, если дальность связи не превышает 45 метров, и не передаются вне данного расстояния).

Предложенный модуль среды моделирования был реализован и опробован. Эксперименты (на основании имитационного компьютерного моделирования) с различными сценариями связи показали, что, сообщения между транспортными средствами могут быть переданы на расстоянии значительно больше 45 метров, а препятствия, влияющие на затухание сигналов, значительно снижают дальность связи.

Заключение

В работе определены требования к моделям сетей и каналов обмена информацией подключенного транспорта, проведен обзор существующих подходов и инструментов для моделирования данных технологий. Также, на основе выбранных требований и инструментов спроектирована архитектура модуля среды моделирования алгоритмов совместного восприятия, позволяющего учитывать задержку передачи сигналов подключенного транспорта.

В качестве решения для дополнения фреймворка OpenCDA предложено использовать сетевой симулятор Artery на базе OMNeT++ и квазидетерминированную геометрическую модель каналов связи GEMV2.

Полученные результаты будут в дальнейшем использоваться для разработки среды моделирования, предназначенной для анализа работы алгоритмов совместного восприятия и для моделирования алгоритма SyncNet.

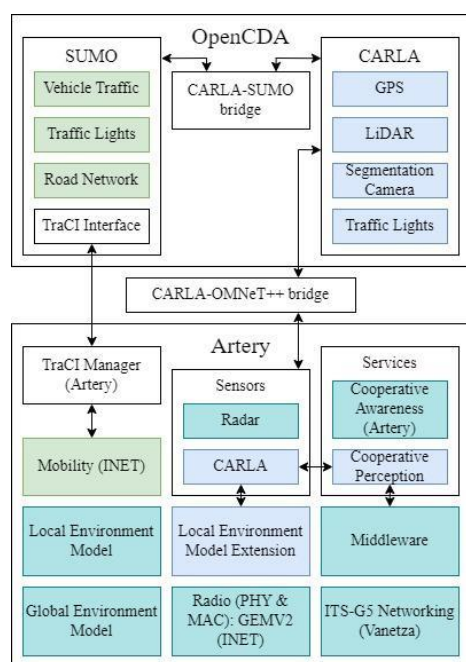


Рис. 1. Интеграция модулей среды моделирования

Список литературы

1. Global status report on road safety 2018. Geneva, 2018.
2. Fagnant D.J., Kockelman K. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations // Transp. Res. Part A Policy Pract. Pergamon, 2015. Vol. 77. P. 167–181.
3. Xu R. et al. OpenCDA: An Open Cooperative Driving Automation Framework Integrated with Co-Simulation // IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. Indianapolis: IEEE, 2021. Vol. 2021-September. P. 1155–1162.
4. Lei Z. et al. Latency-Aware Collaborative Perception // European Conference on Computer Vision 2022. Springer, Cham, 2022. P. 316–332.
5. Boban M., Vinhoza T.T. V. Modeling and Simulation of Vehicular Networks: towards Realistic and Efficient Models // Mobile Ad-Hoc Networks: Applications. IntechOpen, 2011.
6. Lu N. et al. Connected Vehicles: Solutions and Challenges // IEEE Internet Things J. Institute of Electrical and Electronics Engineers Inc., 2014. Vol. 1, № 4. P. 289–299.
7. Mecklenbraüker C.F. et al. Vehicular Channel Characterization and Its Implications for Wireless System Design and Performance // Proc. IEEE. Institute of Electrical and Electronics Engineers Inc., 2011. Vol. 99, № 7. P. 1189–1212.
8. Ur Rehman Khan A., Bilal S.M., Othman M. A performance comparison of open source network simulators for wireless networks // Proceedings - 2012 IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2012. Penang: IEEE Computer Society, 2012. P. 34–38.
9. Weber J.S., Neves M., Ferreto T. VANET simulators: an updated review // J. Brazilian Comput. Soc. Springer Science and Business Media Deutschland GmbH, 2021. Vol. 27, № 1. P. 1–31.
10. Riebl R. et al. Artery: Extending Veins for VANET applications // 2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). Budapest: IEEE, 2015. P. 450–456.
11. Rodriguez-Corbo F.A. et al. Propagation Models in Vehicular Communications // IEEE Access. Institute of Electrical and Electronics Engineers Inc., 2021. Vol. 9. P. 15902–15913.
12. Boban M., Barros J., Tonguz O.K. Geometry-based vehicle-to-vehicle channel modeling for large-scale simulation // IEEE Trans. Veh. Technol. Institute of Electrical and Electronics Engineers Inc., 2014. Vol. 63, № 9. P. 4146–4164.

РАЗРАБОТКА МОДУЛЕЙ СИСТЕМЫ УПРАВЛЕНИЯ И УЗЛОВ МУЛЬТИФУНКЦИОНАЛЬНОГО СТЕНДА ДЛЯ ИССЛЕДОВАНИЯ КРИСТАЛЛИЧЕСКИХ МАТЕРИАЛОВ

Маршутина Е.Н.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе описывается создание модулей системы управления установкой для исследования кристаллических материалов в лабораторных условиях, а также сама конструкция этого многофункционального стенда. Данная разработка выполняется в рамках проекта на базе ФНИЦ «Кристаллография и фотоника» РАН.

Введение

В течение многих лет, с тех пор как первая Нобелевская премия по физике была присуждена В. Рентгену в 1901 г., ученые всего мира исследуют возможности рентгеновского излучения. Это связано с его ценными проникающими свойствами, позволяющими анализировать структуру различных материалов. Рентгеновское излучение является важной частью спектра синхротронного излучения, которое используется на ускорителях заряженных частиц – синхротронах. Задействование их учеными в исследованиях предполагает крайне высокую ответственность, ведь количество синхротронов в мире ограничено, и существует целая очередь в записи на них среди исследовательских групп, а весь исследовательский процесс стоит огромных финансовых вложений. Очевидно, что перед проведением подобных масштабных исследований необходимо тщательно готовиться к эксперименту, моделируя его ход и управляя им сначала в лабораторных условиях.

В центре внимания данной работы находится SCADA программной платформы TANGO (TAsco Next Generation Objects) [1], позволяющей проводить автоматизированные эксперименты и использующейся центрами синхротронного излучения, такими как European Synchrotron Radiation Facility, Soleil, Elettra и прочими [2]. Цель работы состоит в создании программного обеспечения синхротроноподобного стенда для исследования кристаллических материалов и разработке его узлов. Для ее достижения требуется выполнение следующего ряда задач:

- 1) Изучение возможностей, которые предоставляет система TANGO для работы с исследовательским оборудованием;
- 2) Разработка модулей программы для управления шаговыми моторами, детекторами рентгеновского излучения и вспомогательными устройствами в составе установки;
- 3) Проектирование в САПР AutoCad частей стенда, таких как адаптеров для узлов гониометров, держателя линз, вакуумного коллиматора и т.п.

В последующих разделах работы отражена актуальность разработки и раскрыта суть предлагаемого решения с точки зрения программной и конструкторской частей.

Актуальность разработки

На данный момент применение рентгеновских лучей настолько широко, что они используются не только в физике, но и в медицине, химии, геологии и биологии [3–5]. Они составляют основу исследования структуры материалов, что является обязательным этапом в производстве космической электроники, вычислительных механизмов, спортивного инвентаря и т. д. Результаты таких исследований дают существенную информацию о свойствах того или иного материала, могут подтвердить его надежность или, наоборот, обнаружить некоторые дефекты структуры.

В настоящее время невозможно провести полный цикл исследования структуры кристаллического материала в лабораторных условиях, задействуя лишь одну установку. Обычно для этого требуется либо ряд дифрактометров, детекторов и вычислительных узлов, либо синхротрон, что выходит за пределы условий лаборатории [6]. Кроме того, содержание большого количества аппаратуры дорого и сложно реализуемо, а использование синхротрона в своих исследованиях доступно лишь небольшому количеству научных работников. Поэтому крайне необходим единый блок оборудования в виде многофункционального стенда, который мог бы выполнять задачи по исследованию кристаллических материалов. Также необходимо автоматизировать процесс исследования и разработать систему управления, позволяющую дистанционно регулировать параметры в течение эксперимента, исключив избыточный контакт с рентгеновским излучением.

Помимо задач материаловедения, функционал стенда также сможет носить и испытательный характер. Это касается работы элементов, установленных на каналах вывода излучения из синхротрона – специально оборудованных станциях. Перед тем, как любые узлы и оптические элементы попадают в состав синхротронных станций, они подлежат проверке, в связи с чем появляется потребность в тестировочном оборудовании. Им как раз и может послужить разрабатываемый стенд.

Предлагаемое решение

Проектируемая конструкция установки включает в себя несколько башен гониометров, отвечающих за позиционирование в пространстве таких объектов, как монохроматора, линз, образца и детектора. Элементы располагаются на подвижных каретках, которые перемещаются по двум направляющим, закрепленным на оптическом столе высокой жесткости. Управление каретками осуществляется вручную, а позиционирование гониометров планируется выполнять с помощью программы управления. Каждый гониометр имеет свою ось перемещения (например, x , z , ω или χ), между которыми необходимо проектировать переходные детали, т.к. размеры одной по-

движки не всегда совпадают с размерами другой. Между гониометрами также предполагается располагать блоки щелей для придания пучку определенной формы. Сам же пучок будет выделен из источника рентгеновского излучения после прохождения монохроматора: в таком случае он будет иметь необходимую для того или иного опыта длину волны. Вся конструкция будет находиться в защитном боксе с раздвижными стенками, оснащенными возможностью автоматической блокировки. Упрощенная схема описанной установки отражена на рисунке .

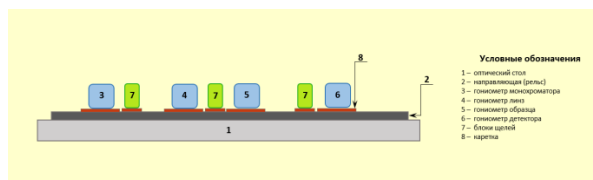


Рис. 1. Схема конструкции проектируемого многофункционального стенда

Программная часть работы базируется на создании отдельных серверов для каждого управляемого устройства в составе стенда, которые впоследствии будут связаны между собой логическим программным блоком. При этом взаимодействие клиента с устройствами будет осуществляться через TANGO протокол.

Рассмотрим концепцию решения на примере сервера для датчика массы образца. Как и любое другое устройство в составе стенда, он принимает на вход команды, которые понятны ему, т. к. вшиты в его процессор. Чтобы можно было управлять датчиком через TANGO, и команды были понятны системе управления, для него нужна промежуточная программа – в среде использования TANGO ее принято называть Device Server. Такой связующий код требуется для каждого устройства, с которого нужно принимать какие-либо данные или отправлять на него команды. Связь между компонентами системы можно схематично представить в виде рисунка , где голубым цветом изображена эта соединительная «прослойка».

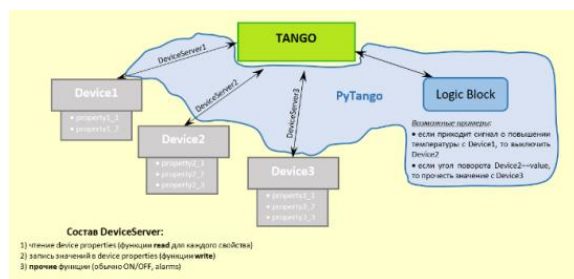


Рис. 2. Схема взаимодействия устройств в составе стенда с системой TANGO

Отметим также, что любой Device Server является по сути процессом-контейнером, внутри которого и создаются соответствующие устройства.

Созданию сервера устройства предшествует написание класса, описывающего его работу, для

чего в настоящей работе используется инструмент генерации базового кода POGO (Program Obviously used to Generate Objects), входящий в средства TANGO [7]. На основе получаемых в нем шаблонов можно добавлять к классу необходимые атрибуты, команды, свойства и функции, что зависит от конкретной задачи. Например, для нашего датчика требуется скалярный атрибут массы и блок считывания данных по COM порту, а для шаговых двигателей понадобятся атрибуты, связанные с углом поворота, и команды позиционирования в пространстве. При этом данные, получаемые с одних устройств, могут отображаться в графическом интерфейсе, транслироваться на другие устройства в качестве входных параметров, или передаваться в логический блок программы, содержащий условия срабатывания тех или иных команд.

Результаты работы

Помимо созданного программного модуля для датчика массы, текущими результатами работы на практике выступают спроектированные конструкции узлов стенда: адаптеров гониометров и держателя рентгеновской линзы. Один из полученных чертежей адаптеров представлен рисунком . Адаптер носит название «Пластина опорная верхняя» и соединяет гониометры перемещения по осям x и y . Для наглядности общий вид башни гониометров вместе с адаптерами отражен на рисунке .

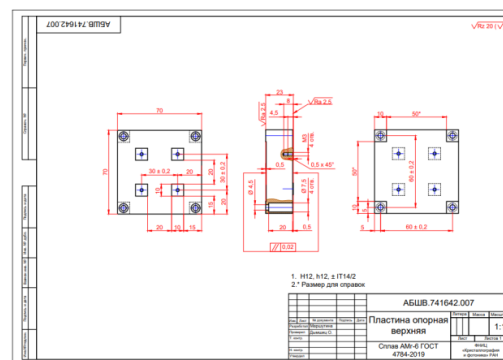


Рис. 3. Пластина опорная верхняя

Конструкция детали учитывает расположение крепежных отверстий на соседних к ней гониометрах и требование к общей высоте узла – центр качения самого верхнего гониометра должен быть выведен на уровень подачи рентгеновского пучка (350 мм). Фрагмент сборочного чертежа держателя рентгеновской линзы показан на рисунке . Конструкция держателя спроектирована так, чтобы надежно фиксировать прижимными элементами тонкий кремниевый чип с рентгеновской линзой в пластине из пластика, которая затем закрепляется в алюминиевом корпусе и монтируется на площадке верхнего гониометра башни. При этом центр чипа с линзами находится на высоте вывода рентгеновского пучка.

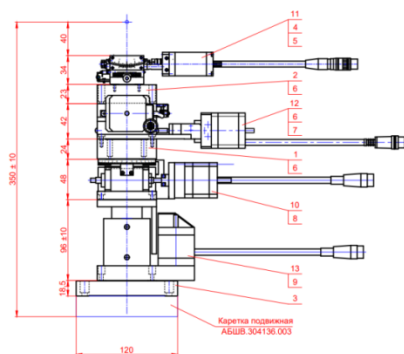


Рис. 4. Общий вид башни гониметров

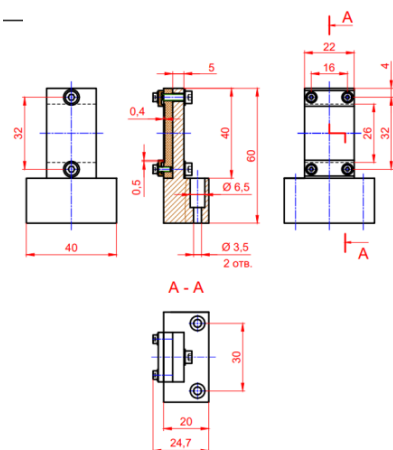


Рис. 5. Держатель рентгеновской линзы

Собранная согласно чертежам башня гониметров представлена на рисунке.



Рис. 6. Узел гониметра рентгеновской линзы

Заключение

Таким образом, разрабатываемый multifunctional stand with a control system is a similar synchrotron research installation, deployable in laboratory conditions. It is expected that this work will serve as a highly practical and economically effective solution in the field of synchrotron research. Thanks to it, it can

be possible not only to study the structure of objects, properties of matter and detect structural defects, but also to conduct tests of synchrotron optics elements for synchrotron radiation stations. Together with the installation, it will allow processing actions during the experiment before its implementation on a real synchrotron, implementing monitoring of the process and controlling its parameters (for example, angles of rotation, distance between the sample and the source of X-ray radiation, etc.).

Список литературы

1. Tango Documentation Contents — Tango Controls 9.3.4 documentation [Электронный ресурс]. URL: <https://tango-controls.readthedocs.io/en/latest/contents.html> (дата обращения: 25.01.2023).
2. History — Tango Controls 9.3.4 documentation [Электронный ресурс]. URL: <https://tango-controls.readthedocs.io/en/latest/overview/history.html> (дата обращения: 25.01.2023).
3. Harlow G.S., Lundgren E., Escudero-Escribano M. Recent advances in surface x-ray diffraction and the potential for determining structure-sensitivity relations in single-crystal electrocatalysis // Curr Opin Electrochem. Elsevier, 2020. – Vol. 23. – P. 162–173.
4. Spence J.C.H. X-ray lasers for structure and dynamics in biology // IUCrJ. International Union of Crystallography, 2018. – Vol. 5. – P. 236–237.
5. Zhang P., Lee Y. il, Zhang J. A review of high-resolution X-ray computed tomography applied to petroleum geology and a case study // Micron. Pergamon, 2019. – Vol. 124. – P. 102702.
6. Кон В.Г. et al. Исследование метода двухволновой рентгеновской дифрактометрии с использованием синхротронного излучения // Кристаллография. Pleiades Publishing Ltd, 2019. – Vol. 64. – No. 1. – P. 29–34.
7. Pogo (Class Generator) — Tango Controls 9.3.4 documentation [Электронный ресурс]. URL: <https://tango-controls.readthedocs.io/en/latest/tools-and-extensions/built-in/pogo/index.html?highlight=pogo> (дата обращения: 25.01.2023).

МОДЕЛИРОВАНИЕ ГЕТЕРОГЕННЫХ ИНФОРМАЦИОННЫХ ВЗАИМОДЕЙСТВИЙ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ

Волгина О.А., Нечаева П.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ,
Институт проблем управления
им. В. А. Трапезникова РАН

Аннотация

Социальные сети являются мощным инструментом, оказывающим влияние на общественное мнение во многих сферах. При этом решение задачи в обла-

сти информационного управления создает основу для изучения характера влияния субъектов реальных социальных сетей друг на друга. В работе описана методика визуализации графа, строящегося на основе топологии, выгруженной из реальной социальной сети. В исследовании представлены результаты численного моделирования поведения сети на основе полученной топологии.

Введение

Социальные сети представляют собой среду со сложной структурой, которая с трудом поддается анализу ввиду динамических изменений. В связи с этим процесс изучения закономерностей распространения информации в реальных социальных сетях является непростой задачей, так как на появление каждой связи влияет большое количество параметров, при этом сеть постоянно растет из-за появления новых пользователей. В таком случае, наиболее приемлемым вариантом становится использование моделей, имитирующих поведение пользователей в социальных сетях [1].

Современная научная литература содержит описания как теоретических, так и численных методов, позволяющих изучать процессы распространения активности в сетях. Имитационные модели получили широкое распространение, так как они позволяют моделировать и анализировать наиболее значимые закономерности динамики изменения мнений в сложных сетях [2].

Любая социальная сеть может быть представлена в виде множества вершин и множества ребер. С помощью вершин в таком графе обозначены участники сети, с помощью ребер – отношения между обозначенными участниками. Большую группу графовых моделей можно разделить на три основных типа.

Графовые модели, дающие количественные оценки, позволяют выявлять основные свойства моделируемых сетей. Выбор конкретной графовой модели зависит от конечной цели исследования и выбранного метода исследования.

Формальное описание модели

Данное исследование основано на описанной ранее модели гетерогенной сети [3].

Динамика системы происходит в дискретном времени t . В каждый момент времени, агент может генерировать информацию только одного типа. Множество узлов задается параметром $N = \{1, \dots, N\}$.

Исходными данными для моделирования сети выступают матрица влияния $R \in \mathbb{R}^{N \times N}$ и количество возможных типов активностей агентов $C = \{c_1, \dots, c_m\}$, где m – количество таких типов.

Параметрами агентов выбраны стохастический вектор интересов $p_i = (p_1, \dots, p_m)$ длины m и порог активации Th . На каждого агента может влиять множество его соседей.

Первоначально производится расчет попарного влияния одного агента на другого согласно матрице влияния. Далее вычисляется суммарное воздействие агентов на одного конкретного за такт. Затем данный

агент выбирает тот тип активации, со стороны которого испытывает наибольшее воздействие. Если сила ни одного из воздействий не превысила установленный порог, то агент на текущем такте не активизируется.

Экспериментальное исследование имитационной модели

Визуализация графов, состоящих из 40 узлов и более, выполнялась с помощью программного инструмента Gephi.

Экспериментальное моделирование проводилось на данных пользователей социальной сети ВКонтакте, выгруженных с помощью API. Для исследования было выбрано два тесно связанных публичных сообщества социальной сети (паблика).

Топология графа строится по следующему алгоритму: выполняется выгрузка идентификаторов пользователей одного паблика и строятся связи между ними; аналогичные действия производятся со вторым пабликом; выполняется поиск связей между подписчиками двух пабликов (пользователями).

Полученная таким образом структурная матрица задаёт ориентированный граф с равнозначными ребрами. Функционал используемого программного приложения Gephi позволяет выполнить укладку графа по силовому алгоритму Force Atlas, позволяющему разделить пользователей сети на кластеры.

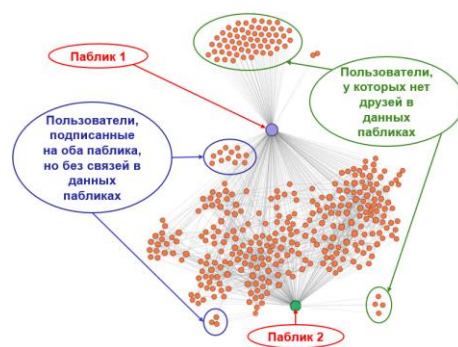


Рис. 1. Укладка графа пользователей социальной сети ВКонтакте по кластерам

Для описанных ниже экспериментов из социальной сети было выгружено 360 узлов: 2 тесно связанных публичных сообщества, 186 подписчиков первого сообщества, 172 подписчика второго сообщества.

В проведенных экспериментах паблики генерировали информацию независимо от пользователей. Под влиянием сообществ их подписчики активировались согласно алгоритму и распространяли активность между собой. Для экспериментального исследования топология графа была зафиксирована и в начальный момент времени пользователи информацию не генерировали.

Генерация мнения пабликами только на начальном такте

Для задания значений порогов активации агентов использовалось треугольное распределение, входными параметрами которого являются интервал и мода распределения. Данная серия экспериментов

проводилась с тремя различными вариантами значений распределения порогов. Активность генерировалась пабликами только в начальный момент времени.

Итоги, полученные в ходе эксперимента, позволяют судить о схожести результатов моделирования топологий на основе моделей случайной генерации графов [4] и моделирования топологии, выгруженной из реальной социальной сети, и подтверждают вывод о том, что увеличение интервала разброса приводит к увеличению числа неактивных агентов.

Генерация мнений пабликами на всех тактах

В данной серии экспериментов значения порогов также задавались согласно треугольному распределению и было выбрано три варианта входных данных. Отличие от предыдущей серии экспериментов заключается в генерации активности пабликами непрерывно на всех тактах.

На основе данной серии экспериментов был сделан вывод о том, что при постоянной генерации активности увеличение порогов слабо влияет на количество неактивных агентов в сети.

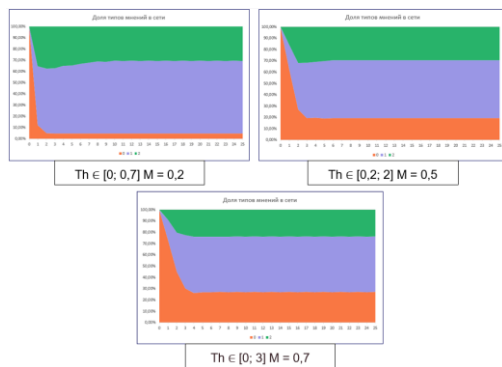


Рис. 2. Распределение интересов агентов по долям с разными порогами активации агентов при постоянной генерации мнений пабликами

Генерация мнений сообществами через заданный период

Следующая серия экспериментов проводилась при различных периодах генерации информации. Паблики генерируют мнения с одинаковым периодом, но второй паблик активируется на полупериод позже первого.

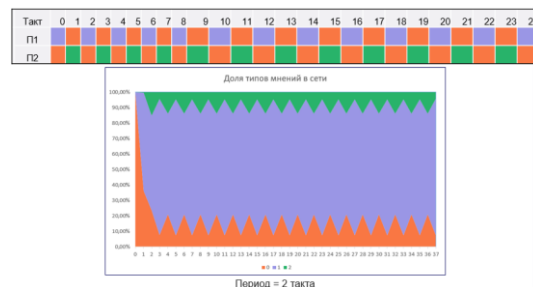
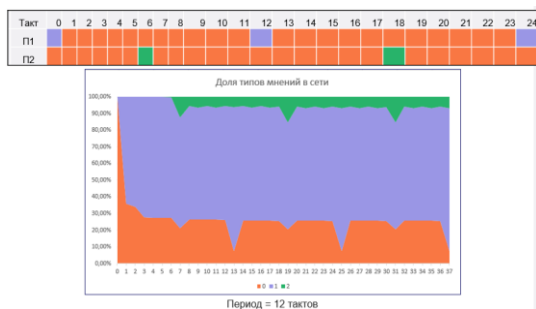


Рис. 3. Распределение интересов агентов по долям при генерации мнений через заданный период

Проведенная серия экспериментов позволила сформулировать следующий вывод: величина периода генерации информации от пабликов слабо влияет на распространение первоначально установившегося преобладающего типа информации по сети. Также важно отметить, что при этом преобладающим типом становится первый сгенерированный тип активности.

Заключение

В данной работе описана апробация выбранной модели на имитационной сети. По результатам апробация прошла успешно, в связи с этим следующим этапом станет сравнение параметров с данными реальной социальной сети. Это даст возможность в дальнейшем указать какой узел является наиболее влиятельным и через какие узлы можно более эффективно распространять информацию по сети. Моделирование на основе агентов, описывающих пользователей сети, позволило создать сетевые модели предпочтительной активности и на их основе исследовать явления распространения мнений при различных сценариях.

Список литературы

1. Бадрызов В. А., Сидельцев В. В. Оценка эффективности распространения информации в социальных сетях с использованием имитационного моделирования // Креативная экономика. – 2018. – Т. 12. – №. 9. – С. 1359-1372.
2. Губанов Д. А., Петров И. В., Чхартишвили А. Г. Многомерная модель динамики мнений в социальных сетях: индексы поляризации // Проблемы управления. – 2020. – №. 3. – С. 26–33.
3. L. Yu. Zhilyakova. Modeling the Structure of MIMO-Agents and Their Interactions / Kuznetsov S., Panov A. (eds) Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science, vol. 1093. Cham: Springer, 2019. P. 3–16. DOI: 10.1007/978-3-030-30763-9.
4. Волгина О. А., Нечаева П. А. Численное моделирование процессов распространения активности в социальных сетях на основе моделирования гетерогенных взаимодействий МИМО-агентов // Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В. Арменского. – 2022.

АЛГОРИТМЫ ИСКУССТВЕННОЙ СЕГМЕНТАЦИИ И ГЕНЕРАЦИИ НАБОРОВ ДАННЫХ ДЛЯ СФЕРЫ РУЧНОГО ПРОИЗВОДСТВА

Козырева Д.Д.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

В данной работе рассматриваются алгоритмы, встроенные в систему автоматической разметки и генерации новых датасетов для обучения нейросетей в сфере ручного производства. Под датасетом понимается набор изображений с разметкой. В работе приведен алгоритм сегментации на основе каскада классических методов компьютерного зрения и его модификация, которая позволила получить больше сегментированных объектов. Также рассматриваются два пути генерации новых изображений и приводятся результаты обучения модели Yolo 5 на сгенерированных данных. Наконец, объяснено использование собственных аугментаций, а не встроенных в Yolo.

Введение

В современном производстве увеличилось количество промышленных задач, которые могут быть автоматизированы с помощью технологий компьютерного зрения. Для их использования зачастую требуются тренировочные наборы данных, разметка которых занимает значительное время. Чтобы оптимизировать этот трудоемкий процесс, была разработана система генерации искусственных наборов данных для обучения нейросетей. В нее были внедрены собственные алгоритмы сегментации и генерации данных, которые будут описаны ниже.

Сегментация

1. Схема 1, каскад классических методов компьютерного зрения и MediaPipe Hands

Разработанный алгоритм сегментации включает в себя четыре стадии отделения объекта от фона, затем полученные изображения пропускаются через фильтрацию и далее используются для генерации новых данных. Подробные характеристики данной схемы можно найти в статье [1].



Рис. 1. Схема каскада классических методов компьютерного зрения и MediaPipe Hands

На первом этапе происходит поиск области интереса с помощью MediaPipe Hands [2] - нейросети по детектированию рук от Google. В результате получаются точки скелета руки. Из полученного списка точек берутся четыре: минимальные и максимальные по осям x, y. От них делается отступ на фиксирован-

ное значение и получается прямоугольник, в который заключена рука с деталью – рассматриваемая область интереса, или ограничивающая рамка.



Рис. 2. Применение MediaPipe к исходной фотографии

Далее выделяется красный канал в прямоугольнике и применяется фильтр Гаусса для размытия границ. По бинарному порогу [3] и порогу Оцу [4] выделяется рука – так получается бинарная маска кисти, используемая на следующих этапах.

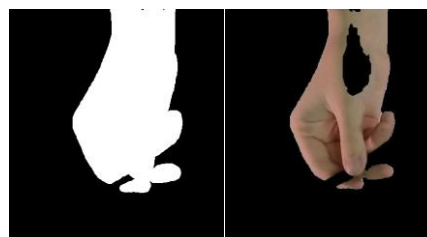


Рис. 3. Рука и ее маска, полученные с помощью цветового порога по красному каналу

Затем с помощью разности кадров [5] рука в области интереса заливается фоновым цветом и в прямоугольнике остается только деталь. Наконец, с помощью цветового порога отделяется деталь и получается ее маска.



Рис. 4. Слева: область интереса исходной фотографии с залитой фоновым цветом рукой. Справа: полученная маска детали

Для данного метода необходима фильтрация, так как на выходе встречаются “бракованные” маски – те, на которых деталь не была выделена корректно. Следующие ситуации были обработаны как некорректные:

1. Деталь была закрыта рукой или же изображение получилось размытым и выделить маску детали не было возможно. Решение: отбрасывать те фотографии, на которых область интереса меньше заданной заранее.

2. Алгоритм нашел много контуров в области интереса, что в большинстве случаев являлось следствием невозможности правильно определить границы детали. Решение: не рассматривать области интереса, в которых найдено больше 3-4 контуров.

2. Схема 2, модификация алгоритма сегментации

В исходном методе использовались ограничивающие рамки, полученные с помощью отступа на заданное значение от крайних точек руки. Данный подход не учитывал возможных размеров деталей (крупный объект мог не войти в область интереса, а для мелкого была бы захвачена область, больше необходимой). Более того, фильтрация, работающая по принципу размерности маски, могла определить как некорректные маски деталей малых размеров. Поэтому, чтобы использовать для обучения нейросетей максимально информативные области интереса и не отфильтровать маски малых деталей, было решено внедрить предобученную нейросеть, детектирующую объект по принципу “деталь – не деталь”. В качестве эксперимента использовалась YOLOv5 [6], обученная на имеющихся размеченных вручную датасетах со втулкой (комплект из 20 различных деталей). Метод применения нейронной сети для выделения области использовался ранее, например, [7], [8]. Более новые модели YOLO не были применены в данном исследовании, так как главное отличие их от пятой версии в производительности, а также более сложной настройке среды выполнения. Цель же эксперимента – проверить, подойдет ли данное решение к поставленной проблеме. Новый алгоритм был проверен на датасете с мясорубкой – комплекте совершенно других деталей. В результате область интереса в несколько раз чаще была найдена в сравнении с предыдущим методом и масок на выходе получилось больше.

Таблица 1. Количество масок, которые были получены на этапе автоматической сегментации (из 1000). Отличие вариантов сегментации в методе поиска области интереса

	MediaPipe Hands + отступ на заданное значение	Обученная на другом комплекте деталей Yolo5
Шнековый вал	53	932
Гайка	44	998
Корпус	11	560
Рукоятка	85	809
Решетка	221	931
Нож	305	919
Винт	1	948



Рис. 5. Области интереса, полученные с помощью, обученной на другом комплекте деталей нейросетью

Генерация

Было реализовано два алгоритма генерации новых изображений – по маскам деталей и по прямоугольным областям интереса. Генерация изображений по прямоугольным маскам была введена, чтобы увеличить информацию о детали, поскольку так маска будет меньше зависеть от качества сегментации.

Генерация изображения происходит следующим образом:

1. Из папки с фонами выбирается одно изображение;
2. Произвольно определяется число деталей на генерируемом изображении в диапазоне от 1 до максимального (параметр, задаваемый пользователем);
3. В цикле по количеству деталей выбирается произвольно деталь и ее маска из папки с сегментированными изображениями, к ним применяются произвольные аугментации и по маске деталь накладывается на фон в произвольной области;
4. Местоположение и размер детали запоминаются и записываются в словарь. По завершении генерации изображения информация из словаря сохраняется в файл разметки.

В алгоритм встроены такие аугментации, как поворот на произвольный угол, инверсия по горизонтальной и по вертикальной осям, увеличение/уменьшение детали. Перерасчет ограничивающей рамки происходит исходя из положений крайних боковых точек маски по осям x , y после поворота, поэтому на выходе получается рамка, плотно прилегающая к границам детали. При применении аугментаций, встроенных в Yolo, перерасчет рамок происходит по принципу заключения повернутого прямоугольника в новый и берутся не крайние точки детали, а крайние точки повернутой области интереса, что приводит к увеличению ограничивающей рамки и включению в нее дополнительной информации, не касающейся детали (рисунок 5).

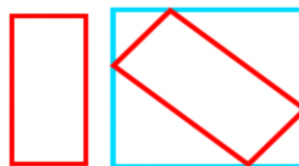


Рис. 6. Изменение ограничивающей рамки при применении аугментации поворота, встроенной в Yolo.

В качестве эксперимента было сгенерировано два набора данных: по маске (dataset_seg) и по области интереса (dataset_rect). Нейросеть обучилась, результаты представлены ниже.

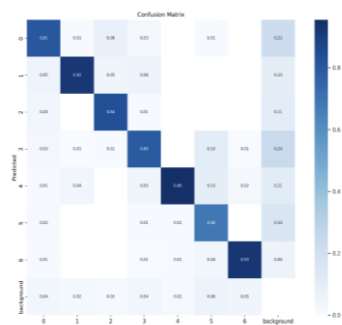


Рис. 7. Confusion matrix, полученная по итогу обучения для dataset_seg

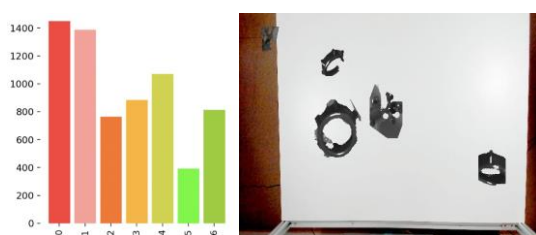


Рис. 8, 9. Распределение классов для dataset_seg и пример сгенерированного изображения из обучающей выборки

Параметры наборов данных:

1. train: 1680 фото мясорубки при ярком освещении;
2. val: 720 фото мясорубки при тусклом освещении (30% выборки);
3. Максимальное количество деталей на фото: 7;
4. Аугментации при генерации: поворот на произвольный угол, инверсия по горизонтали и вертикали, увеличение/уменьшение маски и случайный сдвиг по x, y;
5. Нанесение: по маскам деталей для dataset_seg и по прямоугольным областям интереса для dataset_rect;
6. Фоны: различные фото рабочего места, на котором будет проходить сборка, и градиентный белый фон;
7. IoU: 0.2;
8. Качество изображений: 640 на 480.

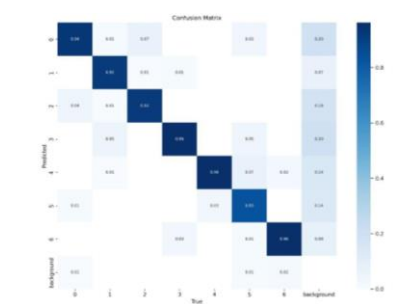


Рис. 10. Confusion matrix, полученная по итогу обучения для dataset_rect



Рис. 11, 12. Распределение классов для dataset_rect и пример сгенерированного изображения из обучающей выборки

Выводы

Наборы фотографий, сгенерированные на основе описанных в исследовании алгоритмов, оказались пригодными для обучения нейросетей и продемонстрировали среднюю точность детектирования 0.89. Модификация алгоритма сегментации позволила в среднем в 8.5 раз увеличить количество масок, получаемых на выходе, а следовательно, и разнообразить генерируемый датасет. Худшим образом алгоритм обрабатывает на деталях малого размера: для датасета по маске точность на детали “нож” составила 0.68, по области интереса – 0.83. Вероятно, детали малого размера стоит фотографировать крупным планом, также, возможно, следует комбинировать методы или использовать лучший по точности. Данные предположения будут исследовано в дальнейшем.

Исследование выполнено с использованием суперкомпьютерного комплекса НИУ ВШЭ.

Список литературы

1. Солдатов А. В., Козырева Д. Д., Мазиков Я.А., Лукашов С. А. Исследование методов отделения объектов от фона на фотографиях // Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов имени Е.В. Арменского. Материалы конференции. - М.: 2022. - С. 116-122.
2. Zhang F. et al. Mediapipe hands: On-device real-time hand tracking //arXiv preprint arXiv:2006.10214. - 2020.
3. OpenCV-Python Tutorials // Open Source Computer Vision URL: https://docs.opencv.org/4.x/d6/d00/tutorial_py_root.html
4. OpenCV-Python Tutorials // Open Source Computer Vision URL: https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html
5. OpenCV: Operations on arrays // Open Source Computer Vision URL: https://docs.opencv.org/3.4.7/d2/de8/group__core__array.html#ga6fef31bc8c4071cbc114a758a2b79c14
6. Jiang P. et al. A Review of Yolo algorithm developments //Procedia Computer Science. - 2022. - vol. 199. - pp. 1066-1073.
7. Lin X., Xu L., Wang Q. Automatic dataset generation for specific object detection //2022 IEEE International Conference on Image Processing (ICIP). - IEEE, 2022. - pp. 3076-3080.
8. Jo H. J., Kim D., Song J. B. Automatic Dataset Generation of Object Detection and Instance Segmentation using Mask R-CNN //The Journal of Korea Robotics Society. - 2019. - vol. 14. - №. 1. - pp. 31-39.

ДЕМОДУЛЯЦИЯ МНОГОПОЗИЦИОННЫХ СИГНАЛОВ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ

Граченко Е.Н.

МИРЭА – Российский
технологический университет,
кафедра радиоэлектронных систем и комплексов
институт радиоэлектроники и информатики

Аннотация

В работе рассматривается эффективность внедрения нейронной сети в квадратурный демодулятор многопозиционных сигналов. Данное исследование проводится в рамках разработки методов и алгоритмов подавления нефлуктуационных помех на основе нейронных сетей (НС).

Введение

В XXI в. широкое распространение получило использование НС в самых разных сферах человеческой жизнедеятельности, таких как экономика, управление производством, медицина, безопасность. С ростом объёмов передаваемой информации возрастает и зашумленность среды распространения информации и, соответственно, потребность в помехоустойчивости. Проводимое исследование рассматривает применение НС в рамках обработки радиотехнических сигналов, в частности многопозиционных сигналов на фоне нефлуктуационных помех. Данная работа проверяет идею, предложенную в [1].

Рассмотрим использование слоя Кохонена в квадратурном демодуляторе многопозиционных сигналов. В первом разделе будет объяснён принцип работы слоя Кохонена. Во втором разделе будут рассмотрены параметры обучения НС, а также сравнение вероятности битовой ошибки с использованием НС и без оной.

«Демодулятор Кохонена»

Каждому j -му нейрону сопоставляется сигнальная точка $W_j = (I_j, Q_j)$ сигнального созвездия, соответствующая определённому символу многопозиционного сигнала. Для входного вектора $x = (I, Q)$ вычисляются евклидовы расстояния $\rho_j(x)$ до точек W_j и «ближайший получает всё» – тот нейрон, для которого это расстояние минимально, выдаёт единицу, а остальные – нули. Следует отметить, что достаточно вычислять линейную функцию сигнала для сравнения расстояний:

$$\rho_j(x)^2 = \|x - W_j\|^2 \quad (1)$$

$$\rho_j(x)^2 = \|W_j\|^2 - 2 \sum_{i=1}^m w_{ji} x_i + \|x\|^2 \quad (2)$$

Здесь $m = 2$, $\|y\|$ – Евклидова длина вектора: $\|y\|^2 = \sum_i y_i^2$.

Задача сводится к поиску номера наибольшего из значений линейных функций:

$$j_{\max} = \arg \max_j \left\{ w_{jI} I + w_{jQ} Q - \frac{1}{2} (I_j^2 + Q_j^2) \right\} \quad (3)$$

Таким образом координаты точки $W_j = (I_j, Q_j)$ совпадают с весами линейного нейрона слоя Кохонена (при этом значение порогового коэффициента $w_{j0} = -(I_j^2 + Q_j^2)/2$).

В результате сигнальное созвездие разбивается на многогранники Вороного-Дирихле, что в условиях равноудалённости сигнальных точек от соседних или центра формирует равномерные области принятия решения. [2]

Таким образом большинство сигнальных созвездий многопозиционных сигналов можно принять за пространство решений нейронной сети.

Результаты моделирования НС

В процессе исследования написаны программы для модуляции сигналов с 4-КАМ, 8-АФМ, 8-ФМ, 16-КАМ и 16-ФМ, прохождения сигнала через радиоканал с шумовой, гармонической и частотно-манипулированной (ЧМ) помехами, демодулятора на основе традиционных методов и на основе слоя Кохонена. В [3] рассмотрены преимущества и недостатки ФМ и КАМ манипуляций.

Для обучения нейронных сетей при параметре скорости обучения равным 1 и обучающей выборке, состоящей из единичных экземпляров каждой сигнальной точки, потребовалось эпох: для 4-КАМ – 1, для 8-АФМ – 2, для 8-ФМ – 180, для 16-КАМ – 5, для 16-ФМ – 150.

В качестве иллюстрации работы НС рассмотрим рис. 1., ОСШ = 7, интенсивность помехи $\mu = 0,9$.

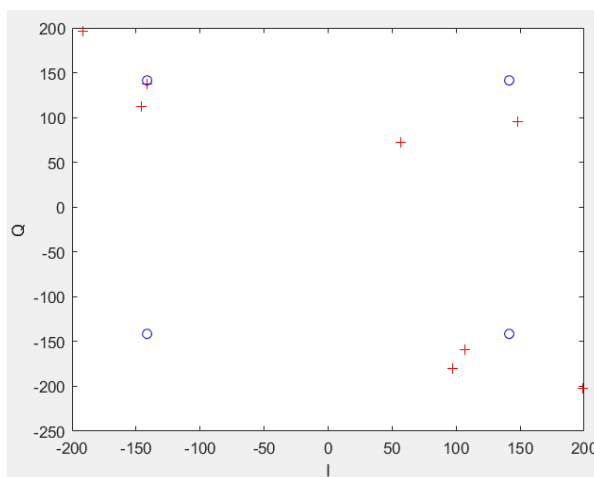


Рис. 1. Пространство решений НС для 4-КАМ для 8 символов (круг – центры кластеров, плюс – входные данные)

Зависимости на рис. 3-4 строились при 100000 символах и интенсивности помехи $\mu = 0,7$, а на рис. 5 при $\mu = 0,1$.

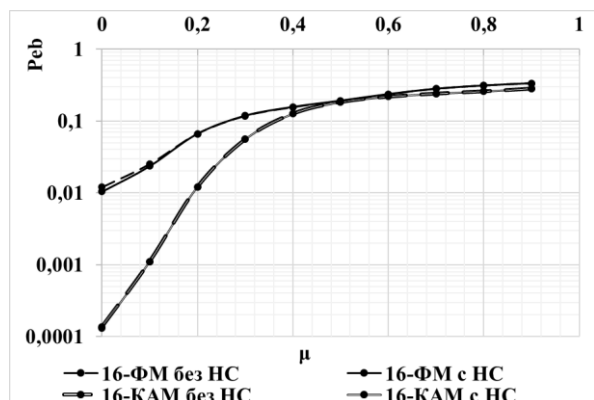


Рис. 2. Зависимость вероятности битовой ошибки от интенсивности гармонической помехи для 16-КАМ и 16-ФМ при ОСШ = 12

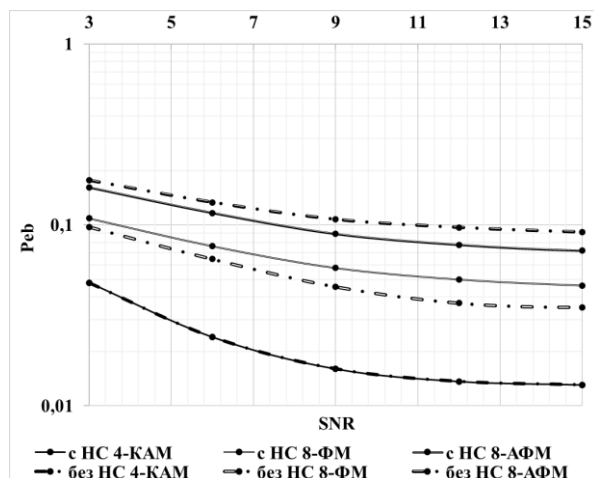


Рис. 3. Зависимость вероятности битовой ошибки от отношения сигнал/шум в присутствии ЧМ помехи для 4-КАМ, 8-ФМ и 8-АФМ

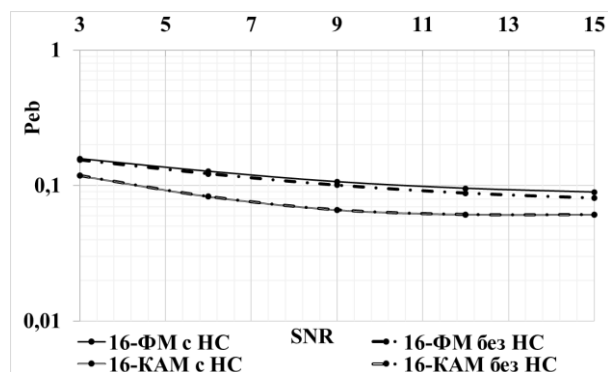


Рис. 4. Зависимость вероятности битовой ошибки от отношения сигнал/шум в присутствии ЧМ помехи для 16-КАМ и 16-ФМ

Проанализировав полученные данные, можно сделать следующие выводы:

- гармоническая помеха оказывает большее воздействие на правильность приёма сигнала, чем ЧМ, что подтверждается в [4-6].

- использование НС в среднем даёт результаты идентичные случаю без использования НС.

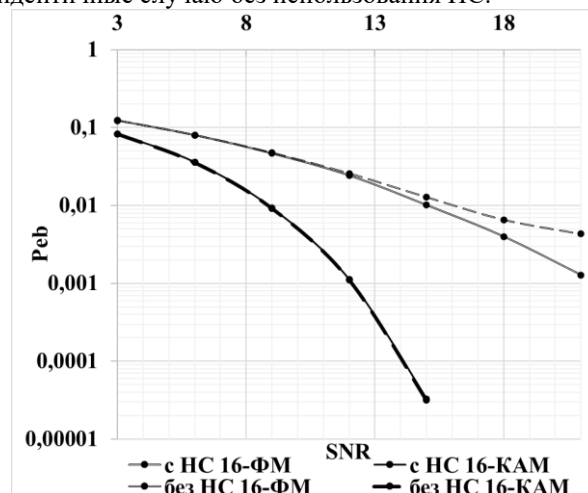


Рис. 5. Зависимость вероятности битовой ошибки от отношения сигнал/шум в присутствии гармонической помехи для 16-КАМ и 16-ФМ

Заключение

Сравнение вероятности битовой ошибки при демодуляции сигналов М-АФМ, М-ФМ и М-КАМ в присутствии гармонической и частотно-манипулированной помех позволяет сделать следующие выводы:

1. При достаточном количестве эпох обучения большинство сигнальных созвездий многопозиционных сигналов можно принять за пространство решений нейронной сети.
2. При отсутствии помех демодулятор с использованием НС обеспечивает нулевую вероятность битовой ошибки.
3. При малых степенях позиционности обучение нейронной сети не требует большого количества эпох, но сигнальное созвездие ФМ требует их на порядок больше, чем КАМ или АФМ.
4. Предварительно можно утверждать, что данный метод демодуляции многопозиционных сигналов имеет как свои преимущества, так и недостатки, но в целом не уступает иным методам демодуляции.

Список литературы

1. Граченко, Е. Н. Возможности применения нейронных сетей для задач фильтрации сигналов с многопозиционной модуляцией от нефлуктуационных помех / Е. Н. Граченко // Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем" ("Радиоинфоком-2022"): Сборник научных статей по материалам VI Международной научно-практической конференции, Москва, 06–10 июня 2022 года. – Москва: МИРЭА - Российский технологический университет, 2022. – С. 172-175.
2. Нейронная сеть Кохонена: [Электронный ресурс] - Режим доступа: https://ru.wikipedia.org/wiki/%D0%9D%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D1%81%D0%B5%D1%82%D1%8C_%D0%

9A%D0%BE%D1%85%D0%BE%D0%BD%D0%B5%D0%BD%D0%B0%D0%98%D0%B4%D0%B5%D1%8F%D0%B8_%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D1%8F, свободный. (Дата обращения: 24.01.2023 г.).

3. Голиков А.М. Модуляция, кодирование и моделирование в телекоммуникационных системах. Теория и практика: учебное пособие / А.М.Голиков. – Томск: Томск. гос. ун-т систем упр. и радиотехники, 2016. – 516 с.

4. Куликов Г.В., Нгуен Ван Зунг, Нестеров А.В., Лелюх А.А. Помехоустойчивость приема сигналов с многопозиционной фазовой манипуляцией в присутствии гармонической помехи // Научные технологии, 2018, № 11. С. 32-38.

5. Куликов Г.В., Нестеров А.В., Лелюх А.А. Помехоустойчивость приема сигналов с квадратурной амплитудной манипуляцией в присутствии гармонической помехи. Журнал радиоэлектроники [электронный журнал]. 2018 № 11 Режим доступа: <http://jre.cplire.ru/jre/nov18/9/text.pdf>. DOI 10.30898/1684-1719.2018.11.9.

6. Анализ помехоустойчивости приема сигналов с квадратурной амплитудной модуляцией на фоне частотно-манипулированной помехи / Г. В. Куликов, А. О. Шамшура, Е. А. Печенин, Е. В. Шаталов // Вестник Воронежского института ФСИИ России. – 2022. – № 2. – С. 9-15.

МЕТОДИКА ПРОГНОЗА ИЗМЕНЕНИЯ ХАРАКТЕРИСТИК БАЗОВОЙ СТАНЦИИ В РЕЗУЛЬТАТЕ МОДЕРНИЗАЦИИ ОБОРУДОВАНИЯ

Валиахметова К.А.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Предлагается методика прогноза изменения количества технических абонентов и нисходящей скорости передачи данных базовой станции (БС) при добавлении новых ячеек для разгрузки существующих слоев оборудования и улучшения качества связи. В частности, исследуются зависимости изменения параметров до и после модернизации базовой станции. Приведены результаты работы двух алгоритмов машинного обучения, их объединения, и проведено их сравнение.

Введение

В настоящее время продолжается рост количества мобильных устройств, и увеличивается доступность мобильной связи. Согласно актуальному отчету компании Ericsson, к концу 2022 году ожидалось, что будет насчитываться около 8,4 млрд абонентов мобильной связи. Прогнозируется, что к концу 2028

года эта цифра увеличится примерно до 9,2 млрд [1]. С ростом подключаемых абонентов к базовой станции возрастает нагрузка на сеть, и это может привести к ухудшению качества связи. Телекоммуникационным компаниям необходимо вовремя проводить модернизацию базовых станций, чтобы обеспечить надежность мобильной сети [2]. Для оптимизации затрат на оборудование необходимо уметь оценивать эффект от разгрузки станции, это поможет предсказывать подходящие модернизации.

С развитием технологий BigData стало возможно получать большие объемы исторических данных работы базовых станций, такие как количество технических абонентов, скорость подключения, нисходящий и восходящий трафик и др. Используя эту информацию, можно получить предсказание изменения технических характеристик БС, используя инструменты машинного обучения [3]. В работе [4] авторы применяют алгоритм прогнозирования временных рядов с целью расчета будущего трафика с LTE-станции ежечасно. Экспериментальные данные позволяют сделать вывод о том, что на основе получаемых прогнозов возможно предвидеть отказ сети и далее определить необходимость модернизаций базовой станции с помощью дополнительного анализа. Но в настоящей работе на этапе обучения модели уже учитываются данные проведенных модернизаций, что не требует дополнительных вычислений в дальнейшем.

Целью данного исследования является получение оптимального по качеству работы алгоритма для предсказания изменения количества технических абонентов базовых станций и скорости нисходящего трафика в результате модернизаций ячеек оборудования с помощью добавления новых слоев. На основании результатов работы алгоритма ожидается получение оценки эффекта разгрузки БС для повышения качества связи.

Описание программной части

Данные для обучения модели содержат основную информацию о модернизациях базовых станций [5]. Для формирования датасета использовались исторические данные конфигурации реальных БС, данные технической статистики с измерениями скорости, трафика и технических абонентов и др. Модернизация проводилась с помощью добавления к станциям слоев 4G-2600, 4G-1800/2100, 4G-800/900. Целевая переменная была сформирована как отношение среднего значения соответствующего параметра за три месяца после модернизации к среднему значению за три месяца до модернизации. Для обучения использовались данные за 15 месяцев, валидация проводилась на 3 последних месяцах.

Для прогноза с использованием кросс-валидации применяются следующие алгоритмы: LightGBM [6], линейная регрессия и сочетание этих моделей.

Для оценки качества алгоритма использовались следующие метрики:

Коэффициент детерминации:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

Средний процент отклонения :

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (2)$$

Симметричный средний процент отклонения:

$$SMAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{\frac{(|y_i| + |\hat{y}_i|)}{2}} \quad (3)$$

Здесь \bar{y} - среднее значение переменной по всем n объектам

Анализ технических характеристик БС

В ходе работы были получены и проанализированы зависимости технических характеристик базовой станции и их влияние на коэффициенты изменения целевых признаков: количества технических абонентов и нисходящей (downlink) скорости в результате модернизаций.

На рисунке 1 изображена зависимость целевой переменной коэффициента изменения количества технических абонентов от среднего за три месяца коэффициента использования базовой станции. Из графика видно, что с увеличением среднего значения нагрузки на БС, коэффициент изменения количества технических абонентов будет снижаться. Таким образом, если коэффициент использования БС (утилизация БС) был высоким, то есть слои станции были нагружены, то с добавлением нового оборудования снижение количества технических абонентов на слоях будет заметнее, так абоненты будут подключаться к новой более свободной ячейке.

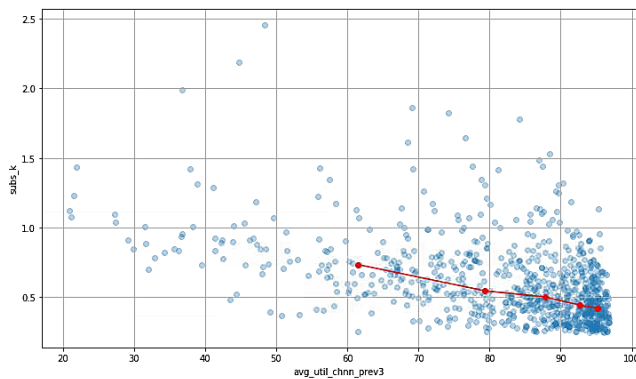


Рис. 1. Зависимость целевой переменной коэффициента изменения количества технических абонентов во время модернизации от среднего коэффициента использования базовой станции до модернизации

На рисунке 2 изображена зависимость коэффициента изменения технической скорости с ростом коэффициента использования БС. В данном случае коэффициент изменения скорости ведёт себя обратно пропорционально коэффициенту изменения количества технических абонентов. Наблюдается рост коэффициента изменения скорости при увеличении средней за три месяца нагрузки на БС. При добавлении нового слоя к сильно нагруженному слою ско-

рость будет расти, так как будет происходить разгрузка ячейки.

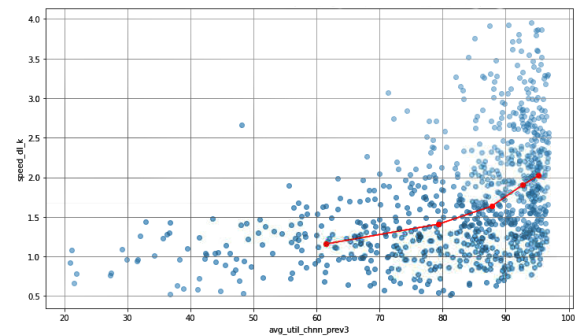


Рис. 2. Зависимость целевой переменной коэффициента изменения технической скорости во время модернизации от среднего коэффициента использования базовой станции до модернизации

На рисунке 3 показан график зависимости коэффициента изменения трафика от технической скорости. Можно заметить, что с увеличением среднего значения скорости подключения за предыдущие три месяца, коэффициент изменения трафика будет снижаться. То есть с добавлением нового слоя часть абонентов подключится к нему, поэтому на исходной ячейке трафик будет снижаться, а скорость расти.

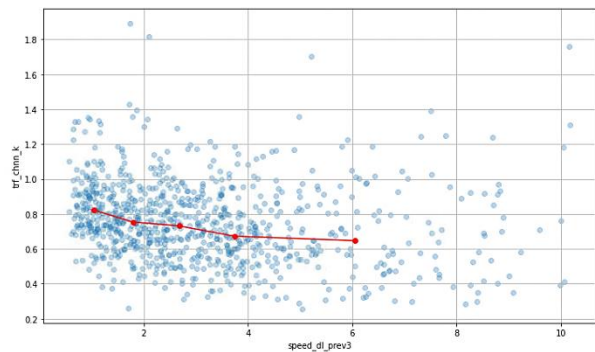


Рис. 3. Зависимость коэффициента изменения трафика во время модернизации от средней технической скорости до модернизации

Результаты работы алгоритма для прогноза количества технических абонентов

В таблице 1 приведены значения метрик по всем примененным алгоритмам машинного обучения для модели изменения количества технических абонентов. Первая строка в таблице — это результат модели LightGBM при обучении на данных без очистки и генерации новых признаков, принятый за бейзлайн. Далее идут метрики для обновленной модели LightGBM, обученной на предобработанных и очищенных от выбросов данных с применением генерации новых признаков. Также был проведен эксперимент с линейной регрессией для исследования наличия линейных зависимостей между признаками. Здесь также использовались очищенные данные. Результат приведен в таблице 1. Коэффициент детерминации составил 0,32.

Так как линейная регрессия показала близкие к LightGBM результаты, был реализован ансамбль этих моделей этих двух моделей. Однако для модели прогноза количества абонентов ансамбль справился хуже, чем обновленная модель LightGBM.

Таким образом, по сравнению с первым бейзлайном коэффициент детерминации увеличился более, чем в два раза, и лучшим по качеству алгоритмом является LightGBM, обученный на преобработанных данных, с коэффициентом детерминации, равным 0,36.

Таблица 1. Метрики моделей предсказания количества технических пользователей базовой станции

Модель	R^2	SMAPE	MAPE
Бейзлайн	0,17	0,34	0,38
LightGBM	0,36	0,31	0,34
Линейная регрессия	0,32	0,31	0,35
Комбинация моделей	0,34	0,31	0,34

Результаты работы алгоритма для прогноза downlink скорости подключения

В таблице 2 приведены значения метрик по всем примененным алгоритмам машинного обучения для модели прогноза downlink скорости подключения. Использовались те же конфигурации моделей, приведенные выше, то есть условия проведения эксперимента сохранились. На бейзлайне был получен результат с коэффициентом детерминации 0,37, а лучшие метрики показала комбинация двух моделей результатов, коэффициент детерминации составил 0,53, SMAPE и MAPE также показали лучшие значения по сравнению с остальными моделями, 0,2 и 0,19, соответственно.

Таблица 2. Метрики моделей предсказания нисходящей скорости подключения

Модель	R^2	SMAPE	MAPE
Бейзлайн	0,37	0,25	0,25
LightGBM	0,5	0,22	0,20
Линейная регрессия	0,51	0,22	0,20
Комбинация моделей	0,53	0,20	0,19

Выигрыш комбинации LightGBM и линейной регрессии может быть связан с тем, что для некоторых признаков наблюдаются более линейные взаимосвязи с целевой переменной, которые находит линейная регрессия. Стоит отметить, что низкие значения коэффициента детерминации могут быть обусловлены сильно зашумленными данными, на которых проводилось обучение

Заключение

В работе получен алгоритм предсказания изменения технических характеристик при модернизации базовых станций. Для прогноза количества технических абонентов лучший результат показала модель LightGBM, комбинация моделей LightGBM и линейной регрессии дают выигрыш в метриках для прогноза нисходящей скорости. Также был проведен анализ параметров характеристик модернизируемых станций. Были установлены качественные зависимости, связанные с разгрузкой базовых станций. К основным гипотезам по улучшению данного исследования и повышению обобщающих способностей моделей машинного обучения можно отнести выделение новых признаков по абонентским характеристикам, исследование вклада соседних базовых станций, то есть признаки ближайших соседей. Полученные результаты могут быть применены мобильными операторами для нахождения оптимальных модернизаций.

Список литературы

1. Ericsson Mobility Report, November 2022 report. [Online]. Available: www.ericsson.com/mobility-report.
2. D. Chmieliauskas and D. Guršnys, "LTE Cell Traffic Grow and Congestion Forecasting," 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2019, pp. 1-5, doi: 10.1109/eStream.2019.8732145.
3. I. A. Tomić, M. S. Davidović and S. M. Bjeković, "On the downlink capacity of LTE cell," 2015 23rd Telecommunications Forum Telfor (TELFOR), Belgrade, Serbia, 2015, pp. 181-185, doi: 10.1109/TELFOR.2015.7377443.
4. A. Y. Nikraves, S. A. Ajila, C. -H. Lung and W. Ding, "Mobile Network Traffic Prediction Using MLP, MLPWD, and SVM," 2016 IEEE International Congress on Big Data (BigData Congress), San Francisco, CA, USA, 2016, pp. 402-409, doi: 10.1109/BigDataCongress.2016.63.
5. Naran Khokshanov, October 23, 2022, "Mobile stations KPI before and after RET optimization", IEEE Dataport, doi: <https://dx.doi.org/10.21227/s8p6-0f81.s>
6. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q. and Liu T.Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 30.

РАЗРАБОТКА УСТРОЙСТВА ДЛЯ САМОКОНТРОЛЯ ЗРЕНИЯ ЧЕЛОВЕКА

Казакова В.Е.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе показана разработка устройства для контроля параметров зрения человека. Обоснована актуальность данного устройства, описан принцип его работы и метод определения патологий зрительного аппарата. Проведено моделирование основного режима работы устройства и одной из версии профилактики болезней. Предложен способ соединения устройства и сервера медицинского учреждения с использованием технологии интернет вещей.

Введение

С появлением и развитием сети Интернет, бытовая жизнь каждого человека резко изменилась, и теперь большинство современных профессий связано с работой с встроенными дисплеями, будь то компьютеры или гаджеты. Известно, что непрерывное использование таких устройств негативно влияет на зрение человека, сказывается в том числе на его остроте и может привести к болезням зрительного аппарата. По данным ВОЗ, на 2019 год в мире насчитывалось более 2.2 миллиардов случаев нарушения зрения, и более миллиарда заболеваний из них являются следствием отсутствия профилактики или лечения [1]. Поэтому отслеживание показателей зрения должно стать обязательной частью жизни каждого человека, который заботится о своем здоровье. Для этого необходимо разработать специальные устройства, облегчающие эту задачу.

Целью данной работы является повышение эффективности проверки зрения в бытовых условиях, т.е. разработка структуры программно-аппаратного комплекса, позволяющего контролировать состояние зрительного аппарата без посещения специализированных клиник. Для ее достижения необходимо изучить зрительную систему человека, выбрать метод измерения параметров зрения и разработать программы для профилактики болезней глаза, в перспективе — направлять полученные данные в медицинское учреждение для отслеживания показателей, назначения профессионального курса лечения и возможной записи на прием к специалисту.

Тезисы включают 5 основных разделов. В первом приводятся аналоги разрабатываемого устройства. Во втором описаны существующие методы проверки зрения и выбран подходящий. Третий раздел посвящен принципу работы разрабатываемого устройства. В четвертом разделе разобран механизм работы модуля для профилактики проблем со зрением. В пятом разделе рассматриваются перспективы разрабатываемого устройства.

Обзор аналогов разрабатываемого устройства

В современном мире проблема автономной проверки зрения достаточно актуальна, поэтому уже существуют устройства, позволяющие это делать. Во многих медицинских учреждениях стоят продвинутое устройства, однако для домашнего использования, людьми без медицинского образования, такие устройства не подходят. Так, например, существуют контактные тонометры [2], позволяющие вне клиники проверять внутриглазное давление. Такой прибор позволяет на ранних стадиях выявить болезни зрительного нерва и развитие глаукомы, а также предоставляет численный результат проверки, интерпретируемый без участия специалиста. Однако такой прибор является дорогостоящим и небезопасным: в основе его работы лежит процесс изменения формы глазного яблока под действием небольшого давления датчика, на основе поведения глазного яблока формируется результат, но неопытный человек при использовании такого устройства может повредить важный орган восприятия, поскольку измерение проходит вплотную к органу.

Другое устройство под названием «Рубин-М» [3] многофункционально и позволяет проверять несколько параметров зрения, от спекл-поля глаза испытуемого до определения его критической частоты слияния мельканий, но в его работе используется лазерное излучение, которое в долгосрочной перспективе может навредить сетчатке глаза, а также вес данного устройства превышает 4 кг, что делает его маломобильным.

Недавно изобретенное устройство от компании EyeQue под названием EyeQue VisionCheck [4] — автоматический быстродействующий авторефрактометр, предназначенный для домашней проверки остроты зрения. Специальное мобильное приложение, идущее в паре к устройству, позволяет быстро измерять и анализировать рефракционную ошибку (параметры, связанные с фокусирующей способностью глаз и выявляющий близорукость, дальнозоркость и астигматизм). Данное устройство более доступно по цене, однако иностранный производитель назначает цену за ежемесячную подписку в долларах, что может стать финансово недоступно пользователям, а также данное устройство не является беспроводным и требует непосредственного подключения к корпусу телефона.

Обоснование метода проверки зрения

Разрабатываемое устройство, наряду с выполнением основной функции, должно так же быть беспроводным, а поэтому — быть соответственно легким и переносимым. Множество методов проверки остроты зрения и здоровья глаза, предлагаемые в специализированных клиниках, в основном не соответствуют этим требованиям. Например, распространенная визометрия не может быть использована в данном устройстве, поскольку требует от устройства больших размеров и нахождения на строго определенном расстоянии от пользователя, а для бытовых условий такие ограничения не отвечают требованиям эргономичности [5]. К тому же, уже суще-

ствующие проекторы знаков мало чем отличаются от бумажной версии визометрии. Многие уже реализованные аппаратно и применяющиеся методы, такие как авторефрактометрия [6], биомикроскопия глаза [7] или офтальмоскопия [8] основаны на совместной работе дорогостоящих линз и требуют присутствие рядом специалиста для непосредственной трактовки результатов, что не соответствует требованию к устройству, предназначенному для самоконтроля. Самым подходящим методом исследования зрения в этом случае является исследование критической частоты слияния мельканий (далее КЧСМ), который будет подробнее рассмотрен в следующем разделе. Преимуществами данного метода является автономность (полученный результат исследования численный и может трактоваться без участия специалиста) и адекватная сложность реализации (для работы устройства не требуются дорогостоящие технологии).

Принцип работы основного модуля

Основная функция модуля — на основе критической частоты слияния мельканий определять степень здоровья зрительной системы человека. Слияние мельканий — процесс, при котором человек перестает различать отдельные мелькания источника света и воспринимает их как непрерывное равномерное свечение. Критическая частота — частота, соответствующая границе перехода мельканий от различимых к слитым. Зрение человека находится в пределах нормы, если данный параметр составляет от 40 до 46 Гц [9]. Если данный параметр выходит за границы нормы, т.е. понижается до 30 Гц и меньше, это может свидетельствовать о развитии патологии в зрительном аппарате человека и необходимости немедленного лечения.

В основе устройства — программируемый контроллер ATmega328, кнопка включения с подтягивающим резистором и светодиод (рис. 1). Этот набор элементов минимально необходим для выполнения устройством основной функции. После включения устройства к питанию однократное нажатие на кнопку запускает процесс измерения КЧСМ: диод начинает мерцать, постепенно увеличивая скорость переключения, на определенной частоте пересекая критическую частоту глаза пользователя. Как только это произойдет, необходимо снова нажать на кнопку, это остановит работу диода и зафиксирует величину измеряемого параметра. После этого устройство снова готово к работе. Алгоритм программы основного модуля представлен на рисунке 2.

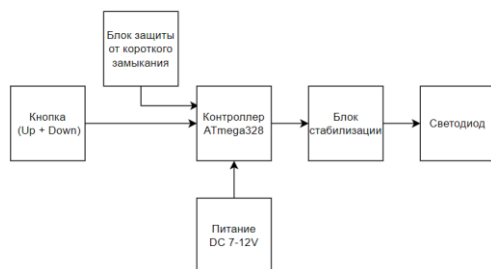


Рис. 1. Структурная схема работы основного модуля

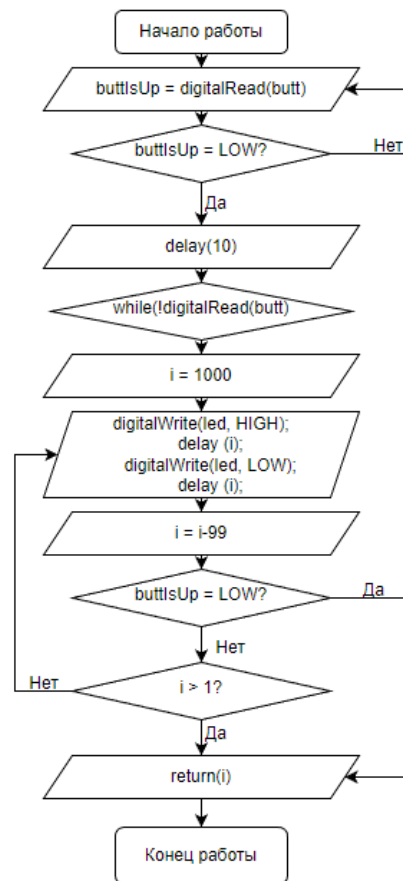


Рис. 2. Алгоритм работы программы основного модуля

Модуль для профилактики болезни глаза.

Данное устройство может использоваться не только для проверки зрения, но и для профилактики различных его заболеваний, а также для тренировки его параметров. Структура таких зарядок является модульной: на наборе из 25 диодов можно организовать множества вариантов зарядок, для этого в микроконтроллер будут подгружены соответствующие варианты программ. На рисунке 3 представлена расширенная структурная схема, содержащая в своем составе дополнительные элементы — микросхемы 74НС595, необходимые для расширения количества выводов микроконтроллера для управления большим количеством светодиодов. На данном наборе компонентов реализована зарядка для глаз, позволяющая тренировать такой параметр зрительной системы как скорость восприятия. Принцип работы таков: после нажатия на кнопку диоды, сформированные в матрицу 5x5, загораются в определенном порядке и показывают некоторое количество цифр с определенной скоростью (количество и скорость показа цифр выбираются пользователем). Чем выше скорость смены цифр, тем сложнее их воспринимать человеческому глазу и обрабатывать их, а при увеличении количества показываемых цифр задача еще усложняется. Алгоритм программы для зарядки представлен на рисунке 4.

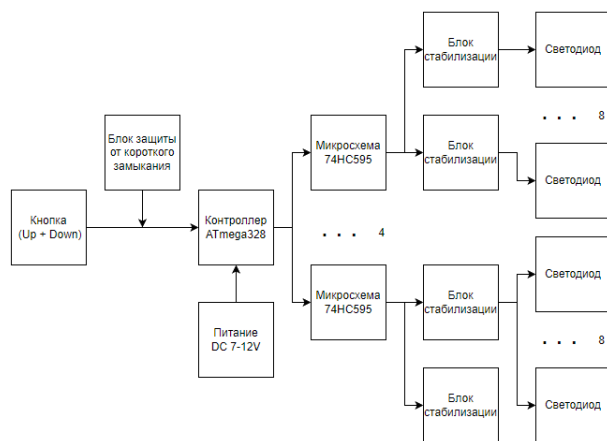


Рис. 3. Структурная схема для тренировок

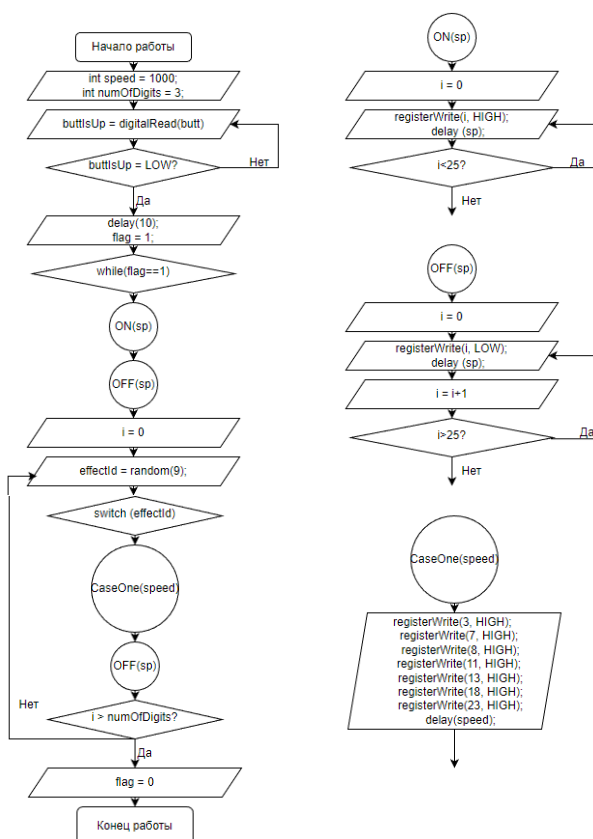


Рис. 4. Алгоритм работы программы для тренировок

Перспективы развития устройства

Для работы с данным устройством, а именно сохранением результатов тестирования, переключением зарядок и редактированием параметров под нужды пользователя будет разработано мобильное приложение. С его помощью устройство должно измерять показатели зрения, обрабатывать данные и формировать отчет (анализ) по результатам контроля. В планах — направлять данные в медицинское учреждение для планирования записи на прием к врачу офтальмологу и назначения курса лечения при первых подозрениях на развитие болезней зрительной системы.

Заключение

Таким образом, в работе были рассмотрены аналоги разрабатываемого устройства, актуальность, рассмотрены методы и выбран подходящий, а также показан принцип его работы и работы модуля тренировки. Планируется дальнейшее развитие устройства, реализация в корпусе и проведение непосредственных тестирований устройства, а также разработка мобильного приложения и подключения к серверам медицинских учреждений.

Список литературы

- ВОЗ публикует первый Всемирный доклад о проблемах зрения [Электронный ресурс] URL: <https://www.who.int/ru/news/item/08-10-2019-who-launches-first-world-report-on-vision> (Дата обращения: 17.01.2023)
- Контактный индикатор внутриглазного давления ТГДц-03 Diathera [Электронный ресурс] URL: <https://medmart.pro/products/tonometr-vnutriglaznogo-davleniya-tgdt-03-diaton> (Дата обращения: 17.01.2023)
- Устройство — приставка «Рубин-М» аппарат полупроводниковый офтальмотерапевтический «красным» спекл-полем [Электронный ресурс] URL: <https://dioptria555.ru/product/rubin-m-apparat-poluprovodnikovyy-ofthalmoterapevticheskij-krasnym-spekl-polem/> (Дата обращения: 17.01.2023)
- УСТРОЙСТВО ДЛЯ ПРОВЕРКИ ЗРЕНИЯ. EYEQUE VISIONCHECK [Электронный ресурс] URL: <https://chipgifts.ru/eyeque-visioncheck> (Дата обращения: 17.01.2023)
- Визометрия — определение остроты зрения [Электронный ресурс] URL: <https://mgkl.ru/uslugi/diagnostika/visometriya-opredelenie-ostroty-zreniya> (Дата обращения: 17.01.2023)
- Аutorefractometry - определение рефракции глаза [Электронный ресурс] URL: <https://mgkl.ru/uslugi/diagnostika/refraktometriya-glaza> (Дата обращения: 17.01.2023)
- Биомикроскопия глаза [Электронный ресурс] URL: <https://mgkl.ru/uslugi/diagnostika/biomikroskopiya-glaza> (Дата обращения: 17.01.2023)
- Офтальмоскопия - осмотр глазного дна [Электронный ресурс] URL: <https://mgkl.ru/uslugi/diagnostika/ophthalmoskopiya-osmotr-glaznogo-dna> (Дата обращения: 17.01.2023)
- Ахмадеев Р.Р., Тимербулатов И.Ф., Кошелев Д.И., Евтушенко Е.М., Тимербулатова М.Ф. Критическая частота слияния мельканий и зрительные вызванные потенциалы при компьютерной нагрузке // Вестник РУДН. Серия: Медицина. 2019. №2. URL: <https://cyberleninka.ru/article/n/kriticheskaya-chastota-sliyaniya-melkaniy-i-zritelnye-vyzvannye-potentsialy-pri-kompyuternoynagruzke> (Дата обращения: 12.04.2023).

ЭНЕРГОЭФФЕКТИВНЫЙ МЕТОД ПЕРЕДАЧИ ДАННЫХ В ГЕТЕРОГЕННОЙ СЕТИ ИНТЕРНЕТА ВЕЩЕЙ ПО СПУТНИКОВЫМ КАНАЛАМ СВЯЗИ

Козлов А.С., Шаповалов М.В.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Работа посвящена разработке метода повышения энергоэффективности при передаче данных по спутниковому каналу связи в рамках глобальной инфраструктуры интернета удаленных вещей. Метод основывается на применении аппаратного шлюза LoRa-Iridium с использованием алгоритма сериализации и сжатия GDEP, а также метода проактивного отслеживания положения спутников.

Введение

Современной тенденцией в сфере «Интернета вещей» (IoT) является изучение внедрения систем сбора данных в удаленных регионах без какой-либо существенной сетевой инфраструктуры. Данная концепция получила название «Интернет удаленных вещей» (IoRT).

Существует ряд областей применения IoRT, в которых использование спутниковой связи (SATCOM) имеет первостепенное значение, поскольку это единственный возможный вариант интеграции локальной сети сбора данных в глобальную сеть [1].

Основным ограничением, которое препятствует интенсивному внедрению спутниковой связи в приложениях IoRT, является высокая стоимость отправки сообщений [2]. Это связано, в первую очередь, с достаточно низкими показателями таких параметров как пропускная способность и энергоэффективность спутниковых каналов связи.

Целью данного исследования является разработка метода повышения энергоэффективности спутниковых каналов связи в сетях IoRT.

Этапы работы:

- а) обзор и анализ методов взаимодействия физических объектов в гетерогенной сети Интернета вещей с использованием спутниковых каналов связи;
- б) разработка модели повышения энергоэффективности спутникового канала связи;
- в) разработка метода повышения энергоэффективности спутникового канала связи;
- г) проведение экспериментального исследования предложенных решений.

Структура IoRT сети с использованием спутникового канала связи

На основе проведенного анализа работ [1], [3], [4], [5] нами была предложена структурная схема гетерогенной сети IoRT с использованием спутникового канала связи, представленная на рисунке 1. Рассматриваемая структура предполагает разделение компонентов сети на наземную и спутниковую со-

ставляющие. Первая необходима для объединения группы удаленных сенсорных устройств в единую сеть для сбора и передачи данных, вторая предназначена для формирования и отправки пакетов сообщений спутниковой группировке, расположенной на низкой околоземной орбите. В качестве связующего элемента между двумя частями выступает граничный шлюз, который принимает данные беспроводной сенсорной сети и подготавливает их для дальнейшей отправки.

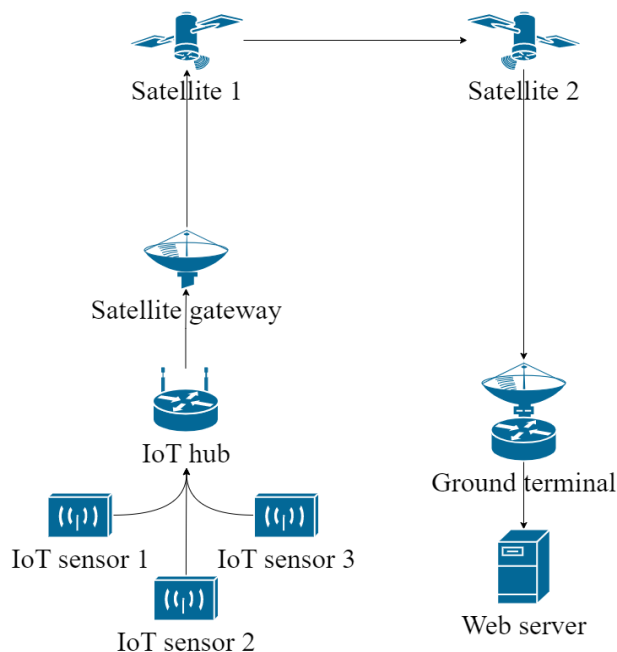


Рис. 1. Структурная схема гетерогенной сети IoRT

В качестве основной беспроводной сети удаленных сенсорных устройств была выбрана реализация LPWAN сети на основе протокола LoRaWAN. Данная сеть обеспечивает наиболее оптимальные показатели с точки зрения энергоэффективности и дальности действия.

В качестве поставщика спутниковой связи была выбрана спутниковая система Iridium, которая обеспечивает покрытие в труднодоступных районах крайнего севера и Арктики.

Модель повышения энергоэффективности

В рамках проводимого исследования нами разработана комплексная модель повышения энергоэффективности (рис. 2). Она служит основой при внедрении и отработке энергоэффективного метода передачи данных по спутниковому каналу связи Iridium. Модель содержит теоретическое и математическое описание внедряемых решений, главные из которых это использование алгоритма сериализации и сжатия данных GDEP [3, 6], а также применение проактивного отслеживания положения спутников [7].

Модель позволяет исключить попытки передачи сжатых сообщений в момент, когда спутники не находятся в оптимальном положении на орбите относительно наблюдателя на Земле.

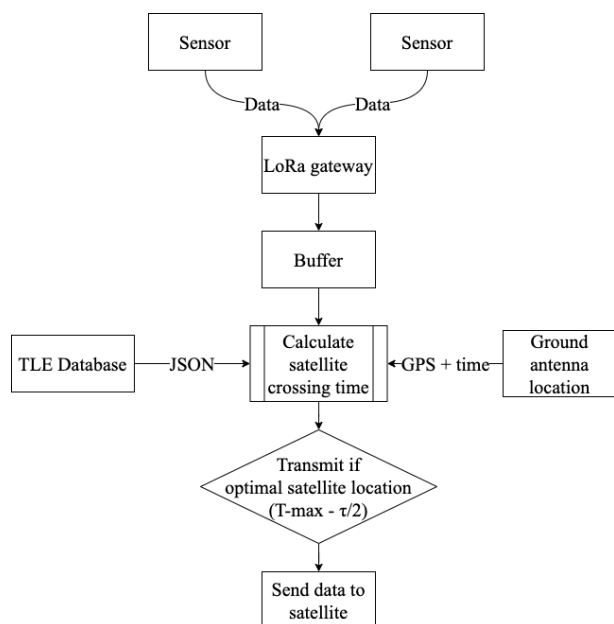


Рис. 2. Модель повышения энергоэффективности

Метод повышения энергоэффективности

В основе метода лежит алгоритм, позволяющий заранее определять временные окна, в момент которых пролетающие спутники имеют оптимальные параметры движения по орбите.

В качестве основного параметра для оценки оптимальности положения спутников был выбран параметр угла места или угол возвышения над локальным горизонтом.

Работа предложенного алгоритма строится в следующем порядке:

- 1) Сперва составляется расписание пролета спутников, которые удовлетворяют заранее выбранному углу места. В качестве исходных данных используются физические параметры движения спутников по орбите, представленные в специальном двустрочном формате TLE;

- 2) Далее запускается основной цикл работы, в процессе которого сравнивается локальное время с расписанием;

- 3) В момент наступления события по расписанию, инициируется проверка буфера на наличие сообщений для отправки;

- 4) При наличии сообщений они записываются в буфер передающего модуля антенны и далее отправляются на спутники.

В дополнение к алгоритму проактивного отслеживания, в методе используются 3-х ступенчатый алгоритм сериализации и сжатия GDEP [3]. С помощью него происходит последовательное сжатие передаваемых данных с использованием протокола Protobuf, а также кодирование на основе метода Хаффмана.

Рассмотренный этап обработки данных осуществляется на граничном шлюзе сети LoRa – Iridium.

Экспериментальная установка

Для экспериментального исследования предложенного метода используется программно-аппаратная установка, схема которой представлена на рисунке 3.

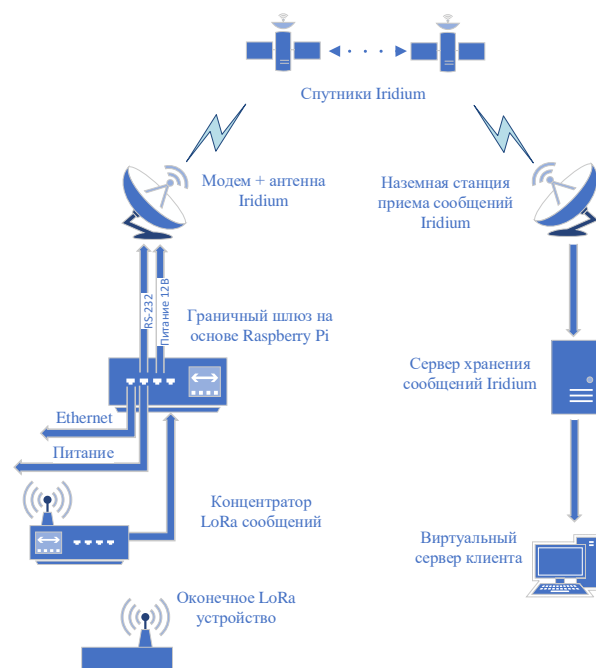


Рис. 3. Схема экспериментальной установки

Алгоритм работы макета:

- 1) Оконечные LoRa – устройства, в виде беспроводных модулей Heltec Wireless Stick формируют сообщения, которые эмулируют данные с датчиков.

- 2) Сгенерированные сообщения поступают на модуль концентратора LoRa, выполненного на основе платы IC880a. Далее полученные данные поступают на шлюз LoRaWAN – Iridium в формате JSON.

- 3) Полученные шлюзом LoRaWAN – Iridium сообщения сохраняются на локальном MQTT сервере. Шлюз реализован на основе одноплатного компьютера raspberry pi 3.

- 4) В соответствии с составленным расписанием метода проактивного отслеживания положения спутников сообщения извлекаются и упаковываются в SBD пакеты.

- 5) Отправка сжатых сообщений на спутники с помощью терминалов спутниковой связи Iridium на базе модемов 9523, 9602, 9770 и антенны Hiirshman ITAS.

- 6) Запись на стороне шлюза, а также на стороне сервера принимающей наземной станции основных характеристик, относящихся к процессу передачи данных — статус отправки, размер исходного сообщения JSON, задержка при отправке сообщений.

- 7) Получение отправленных сообщений на выделенном виртуальном сервере хостинг-сервиса Ruweb.

Результаты эксперимента

Эксперимент состоял из двух этапов. На первом этапе осуществлялась отправка пакетов данных без использования алгоритма проактивного отслеживания спутников. Второй этап проводился уже с использованием данного алгоритма.

По результатам проведенного нами эксперимента были получены данные о параметрах задержки при отправке сообщений на спутники.

Использование алгоритма проактивного отслеживания положения спутников позволило сократить среднее время задержки отправления сообщений в 2,6 раза.

Необходимо отметить, что время задержки характеризует время активного состояния передатчика, в процессе которого происходит повышенное энергопотребление [8].

Таким образом, сокращение времени задержки, полученное в результате эксперимента, позволило повысить энергоэффективность передачи данных и увеличит время автономной работы передатчика в 2,6 раза.

С учетом того, что в эксперименте не использовался алгоритм сериализации и сжатия GDEP можно утверждать, что совместное использование метода проактивного отслеживания положения спутников и метода GDEP позволяет повысить энергоэффективность передачи данных по спутниковому каналу связи более чем в 20 раз.

Дальнейшее развитие

С целью дальнейшего совершенствования предложенного метода планируется реализовать следующие мероприятия:

1) Внедрение системы планирования отправки сообщений на основе приоритетов. Это связано с тем, что в очереди на передачу могут быть сообщения, которые необходимо отправить в момент наступления определенного события (например сообщение о начавшемся пожаре). Для этого предполагается установить пять классов приоритетов: “критическая ошибка”, “ошибка”, “предупреждение”, “информация”, “отладочное сообщение”. Для каждого вида сообщения из перечисленных планируется реализовать отдельное расписание отправки.

2) Реализация системы на основе спутниковых точек доступа Iridium-go для получения не только автономности от сети сотовой связи, но и возможности работы от аккумулятора в течение продолжительного времени.

Заключение

Предложенный в работе метод позволит увеличить энергоэффективность передачи данных в гетерогенных сетях интернета вещей со спутниковым каналом связи более чем в 20 раз и, как результат, будет способствовать более массовому распространению решений «интернета удаленных вещей» с использованием спутниковых каналов связи.

Список литературы

1. Ivan Lysogor, Voskov L., Rolich A., Efremov S. G. Study of Data Transfer in a Heterogeneous LoRa-Satellite Network for the Internet of Remote Things // Sensors. 2019. Vol. 19. No. 15. P. 1-17.
2. Liu, Z.; Li, J.; Wang, Y.; Li, X.; Chen, S. HGL: A hybrid global-local load balancing routing scheme for the Internet of Things through satellite networks. Int. J. Distrib. Sens. Netwo. 2017, 13.
3. Voskov L., Rolich A., Bakanov Gleb, Podkopaeva Polina. Gateway Data Encoding, Packaging and Compression method for heterogeneous IoT-satellite network, in: 2021 XVII International Symposium "Problems of Redundancy in Information and Control Systems" (REDUNDANCY). IEEE, 2021. doi P. 34-38.
4. Смусев М. А., Хайрулин Н. Р. Энергоэффективный метод взаимодействия в гетерогенной сети интернета вещей с использованием спутниковой связи : специальность 11.04.02 «Инфокоммуникационные технологии и системы связи» : диссертация магистра/ Смусев Михаил Алексеевич, Хайрулин Никита Рустамович; Национальный исследовательский университет «Высшая школа экономики» - Москва, 2022. - 66 с. - Библиогр.: с. 60-66. - Место защиты: Московский институт электроники и математики. - Текст : непосредственный.
5. Lysogor I., Voskov L., Rolich A., Efremov S. G. Energy efficient method of data transmission in a heterogeneous network of the Internet of things for remote areas, in: 2019 International Siberian Conference on Control and Communications (SIBCON). Proceedings. Tomsk : Tomsk State University of Control Systems and Radioelectronics (TUSUR), 2019. P. 1-6.
6. Национальный исследовательский университет «Высшая школа экономики»//Выпускные квалификационные работы студентов НИУ ВШЭ// Лысого И.И. Исследование и разработка энергоэффективного метода взаимодействия в гетерогенной сети Интернета вещей с использованием спутниковой связи. [Электронный ресурс]. Режим доступа: <https://www.hse.ru/edu/vkr/219504583>, (дата обращения 15.12.2022).
7. Дубровин, М. Г. Концепция проактивного мониторинга и управления объектами ИТ-инфраструктуры / М. Г. Дубровин. - Текст : электронный // ИТНОУ. - 2020. - № 1. - С. 44-49. <https://cyberleninka.ru/article/n/kontseptsiya-proaktivnogo-monitoringa-i-upravleniya-obektami-it-infrastruktury/viewer> (дата обращения: 13.01.2023). - Режим доступа: Научная электронная библиотека «КиберЛенинка».
8. Gomez, Darroudi S. M., Naranjo H., Paradells J. On the Energy Performance of Iridium Satellite IoT Technology // Sensors. 2021. Vol. 21. No. 7235. P. 1-14.

РАЗРАБОТКА ВЕБ-ИНТЕРФЕЙСА ДЛЯ СТЕНДА УДАЛЕННОЙ РАБОТЫ С ОБОРУДОВАНИЕМ

*Ан Д.А., Щербаченко Н.А., Будаев М.В.,
Карманова Н.Д., Мосолов А.М.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе рассматривается создание программного обеспечения для подключения к удаленному стенду для работы с ПЛИС. В ходе выполнения работы был разработан веб-интерфейс, предоставляющий пользователю возможность работать с ПЛИС [1] De10-Lite удаленно, система авторизации, сценарии работы, система передачи команд управления от веб-сервера на подключение оборудования, асинхронная обратная связь с пользователем.

Введение

Выполнение лабораторных работ, требующих наличие специального технического оснащения – особенность обучения в инженерных высших учебных заведениях нашей страны. Эта особенность является как преимуществом, так и недостатком образовательных программ. С одной стороны, ученик имеет возможность напрямую взаимодействовать с тем или иным оборудованием, получая тем самым новый опыт. С другой стороны, возникает ряд проблем. Например, студент не всегда имеет возможность находиться в учебном заведении. Ярким примером тому служит прошедшая пандемия Covid-19. Помимо этого, не всегда оборудования хватает на всех студентов, и за выделенное на дисциплину время не все могут достаточно времени уделить выполнению работ.

Главной целью работы является создание веб-интерфейса для стенда взаимодействия с ПЛИС (рис 1). Это позволит студентам МИЭМ НИУ ВШЭ выполнять лабораторные работы, требующие взаимодействие с ПЛИС, с использованием любого устройства, имеющего выход в Интернет, без необходимости присутствия в университете. Так же, это позволит ученикам выполнять работы в любое время, вне зависимости от наличия заданий по дисциплине, для которой требуется выполнение этих работ.

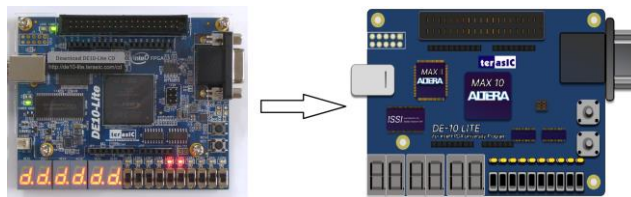


Рис. 1. ПЛИС DE10-Lite физическая (слева) и графическая (справа)

Алгоритм работы

Алгоритм работы системы представляет собой взаимодействие пользователя с разработанным веб-интерфейсом, которое позволяет передавать команды управления на ПЛИС De10-Lite.

Перед описанием алгоритма работы, следует сказать, что стенд работает в асинхронном режиме – то есть, ответ на действия пользователя приходит не сразу, а через некоторое время на электронную почту. Это сделано для того, чтобы несколько пользователей могли взаимодействовать с одним стендом одновременно.

Рассмотрим алгоритм работы для одного пользователя. Веб-приложение размещено на физическом сервере УЛ САПР МИЭМ НИУ ВШЭ [3], доступ к которому осуществляется после авторизации пользователя посредством “Единого личного кабинета НИУ ВШЭ”. К серверу подключен удаленный стенд, представляющий собой несколько плат ПЛИС и видеокамер, отслеживающих поведение плат при отладке кода, подключенных между собой. Пользователь, используя веб-интерфейс, представленный на рисунке 2, загружает файл с сценарием выполнения кода на ПЛИС. Сценарий отладки должен быть представлен на алгоритмическом языке Псевдокод. Стоит отметить, что каждый файл после загрузки проходит обязательную валидацию, которая позволяет уменьшить число ошибок компиляции кода. В случае успешной загрузки файла со сценарием, файл приходит на сервер, откуда он загружается на ПЛИС и компилируется. В это время, установленная в стенд и направленная на ПЛИС, записывает видео и отправляет его на сервер, где происходит его анализ с помощью библиотеки OpenCV [2]. В результате анализа формируется ответ, содержащий информацию о всех изменениях, происходящих с ПЛИС: загорании диодов и семисегментных индикаторов. Ответ отправляется на веб-интерфейс, благодаря чему в динамическом режиме меняется состояние веб-компонентов, происходит изменение их отображения в приложении. Также, видеозапись, сделанная внутри стенда, приходит пользователю на электронную почту.

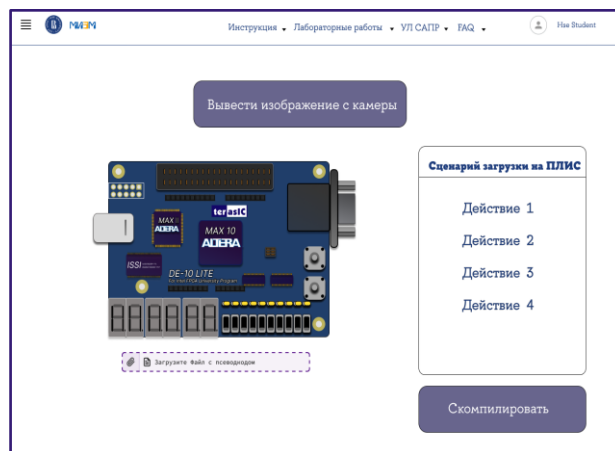


Рис. 2. Главная страница веб-интерфейса



Рис. 3. Вид серверной ячейки стенда

Стек технологий

В работе используется стандартный стек технологий современной веб-разработки. А именно, для проектирования шаблона веб-интерфейса был выбран UI/UX редактор Figma [4], также при работе был задействован официальный стиль НИУ ВШЭ для веб приложений, включающий в себя 8-колоночную систему сетки с соответствующим набором цветов и шрифтов. В качестве основного инструмента для реализации клиентской части приложения в работе используется Javascript библиотека React [5] с рядом дополнительных утилит, таких как: Redux Toolkit, Typescript, Vite, Ку. В свою очередь сервер приложения был запущен на платформе Nginx [6], что позволяет решить проблему “Голодания потоков” сервера в самом начале разработки. Помимо этого, данный выбор обусловлен тем что, имея куда меньшую аппаратную мощность сервера на Nginx, можно получить ту же производительность, в сравнении с обычным сервером на Apache. Также, исходя из выбора развертывания физического сервера на Nginx в работе в качестве платформы для Back-end разработки был выбран Node.js [7] с фреймворком Express. Это объясняется тем, что Nginx использует подход монопоточности как и в Node.js. На сервере размещается программное обеспечение, принимающее поток с камеры посредством OBS и анализируют изображение с помощью библиотеки OpenCV. Алгоритмы компьютерного зрения проверяют, какие диоды и семисегментные элементы загорелись, после чего сервер отправляет ответ на веб-интерфейс.

Для поддержания работоспособности и непрерывной отладки всей разработанной системы используется специальное ПО Docker. Которое позволяет реализовать принцип CI/CD и подгружать зависимости в любой Unix подобной системе, упакованные в Docker-контейнер.

Заключение

Данная работа призвана сделать возможным выполнения в режиме онлайн лабораторных работ МИЭМ НИУ ВШЭ, связанных с программированием на ПЛИС, а также давать полное понимание того,

что происходит с платой, как если бы работа выполнялась очно. Это позволит сократить расходы на оборудование благодаря распределению нагрузки на каждую ПЛИС во избежание перегрева платы и последующего износа. Также, у студентов появится возможность выполнять работы из дома, что снизит нагрузку на преподавателей, и ускорит выполнение лабораторных работ студентами.

В результате проведенной работы был разработан полностью готовый веб-интерфейс для удаленной работы с ПЛИС. Разработанная система авторизации, система передачи команд управления от веб-сервера на подключение оборудования, написаны сценарии работы со стендом.

Список литературы

1. Клайв Максфилд. Проектирование на ПЛИС. Архитектура, средства и методы. Курс молодого бойца. Изд-во Додэжа XXI, 2007. 408 с.
2. Adrian Kaehler, Gary Bradsky. Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library. Изд-во O'REILLY, 2018. 369 с.
3. Учебная лаборатория систем автоматизированного проектирования [Электронный ресурс]. Режим доступа: <https://miem.hse.ru/edu/ce/cadsystem/>. - Дата доступа: 26.01.2023.
4. Figma [Электронный ресурс]. Режим доступа: <https://www.figma.com/>. - Дата доступа: 26.01.2023.
5. Mark Tielens T. React in action. Изд-во Manning, 2019, 368 с.
6. Димитрий Айвалиотис. Администрирование сервера NGINX. Изд-во ДМК Пресс, 2018. 288 с.
7. Хэррон Дэвид. Node.js. Разработка серверных веб-приложений на JavaScript. Изд-во ДМК Пресс, 2012, 144 с.

ПОДГОТОВКА ОБРАЗА, ТЕСТИРОВАНИЕ И РАЗРАБОТКА УТИЛИТЫ ДЛЯ ОДНОПЛАТНОГО КОМПЬЮТЕРА RASPBERRY PI 4 НА БАЗЕ ОС ALT LINUX

*Ранасингхе М.П.К.К., Кантута З.Д.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Данная статья посвящена разработке специализированного дистрибутива на базе операционной системы (далее ОС) ALT Linux с набором необходимых инструментов для взаимодействия с интерфейсами одноплатного компьютера семейства Raspberry Pi 4. Подготовка осуществлялась на основе базового образа с добавлением специализированной авторской утилиты, позволяющей упростить первоначальное конфигурирование и управление интерфейсными пинами GPIO на Raspberry Pi, через бэкенд REST

API путем выполнения HTTP-запросов к конечным точкам /pins и /pins/<id>.

Для существенного сокращения времени на создание конечного решения с принципом обеспечения целостности и дальнейшей поддержки продукта использован инструмент Sisyphus (Сизиф) от проекта ALT Linux Team.

Введение

Идея разработки данного образа заключается в необходимости, с одной стороны, поддержать развитие отечественного программного обеспечения (ПО), что является первоочередной задачей для развития страны, промышленности и внутреннего потенциала в области электроники не зависящего от монопольных иностранных продуктов. С другой, позволит специалистам в сфере развития систем машиностроения и интернета вещей облегчить тестирование и разработку желаемых конечных продуктов перенести основное внимание непосредственно на разработку, а не заниматься настройкой и оптимизацией ПО для взаимодействия с одноплатными компьютерами.

Кроме того, данное решение является уникальным для управления одноплатным компьютером Raspberry Pi 4, а размещение образа в общедоступном репозитории Sisyphus [1,2], даст возможность другим инженерам в данной области использовать текущее базовое решение, как основа при разработке тренажеров или иных образов для аналогичных платформ с разной архитектурой.

Критерии, на которые уделялось особое внимание при создании образа, было использование максимально возможного набора библиотек для взаимодействия с интерфейсами, удобное пользовательское меню, быстроедействие для текущей архитектуры и минимальный размер исходного образа, что подтверждается проведенными тестами.

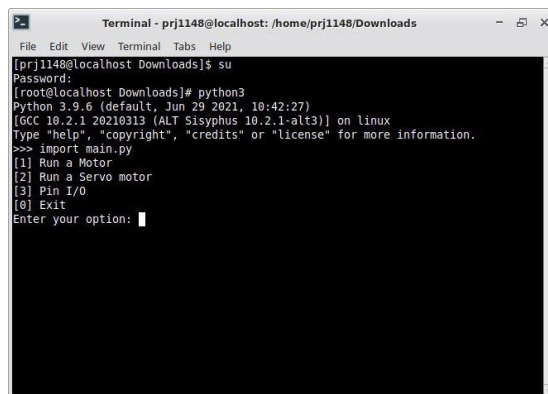
Основная часть

Для взаимодействия инженера с периферийными устройствами и конфигурирования необходимых выходов, для управления подключенных модулей и датчиков к одноплатному компьютеру, внутрь образа была интегрирована консольная утилита, написанная на языке Python, для управления инициализации функций взаимодействия с тем или иным подключенным устройством.

Запуск основного скрипта осуществляется с помощью интерпретатора Python с указанием параметра main.py путь, к которому по умолчанию прописан в глобальной переменной окружения.

После запуска, оператору показано меню и предоставлена возможность выбора необходимых опций, тем самым устанавливая различные предопределенные.

Главное меню изображено на рис. 1.



```
Terminal - prj1148@localhost: /home/prj1148/Downloads
File Edit View Terminal Tabs Help
[prj1148@localhost Downloads]$ su
Password:
[root@localhost Downloads]# python3
Python 3.9.6 (default, Jun 29 2021, 10:42:27)
[GCC 10.2.1 20210313 (ALT Sisyphus 10.2.1-alt3)] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import main.py
(1) Run a Motor
(2) Run a Servo motor
(3) Pin I/O
(0) Exit
Enter your option: 
```

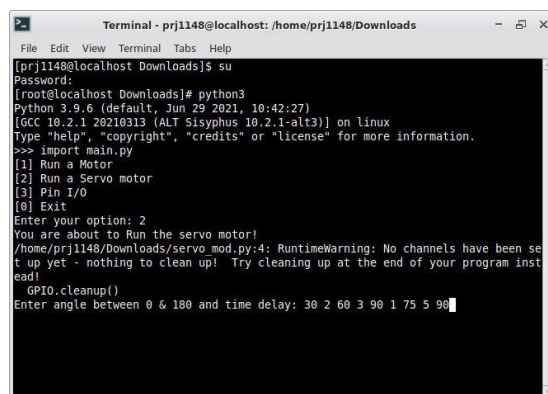
Рис. 1. Главное меню

Рассмотрим настройку одноплатного компьютера на примере подключения к его интерфейсам серводвигателя.

Управление серводвигателями Raspberry Pi очень простое, и благодаря небольшому размеру и весу, используются во многих областях.

В отличие от шаговых двигателей серводвигателями можно управлять с помощью одного GPIO.

После запуска скрипта main.py оператору в главном меню необходимо выбрать пункт 2 - управление серводвигателем (рис. 2).



```
Terminal - prj1148@localhost: /home/prj1148/Downloads
File Edit View Terminal Tabs Help
[prj1148@localhost Downloads]$ su
Password:
[root@localhost Downloads]# python3
Python 3.9.6 (default, Jun 29 2021, 10:42:27)
[GCC 10.2.1 20210313 (ALT Sisyphus 10.2.1-alt3)] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import main.py
(1) Run a Motor
(2) Run a Servo motor
(3) Pin I/O
(0) Exit
Enter your option: 2
You are about to Run the servo motor!
/home/prj1148/Downloads/servo_mod.py:4: RuntimeWarning: No channels have been set up yet - nothing to clean up! Try cleaning up at the end of your program instead!
GPIO.cleanup()
Enter angle between 0 & 180 and time delay: 30 2 60 3 90 1 75 5 180
```

Рис. 2. Запуск серводвигателя

Далее, предложено ввести любой угол в диапазоне 0 до 180 и время задержки для следующего угла (например: 30 2 60 3 90 1 75 5 180).

После, создаётся и выполняется сценарии с указанными ранее параметрами, а по завершению - возврат в главное меню.

Стоит отметить, что в проекте так же реализованы механизмы обратной связи, обеспечивающие обратную связь состояния каждого тестируемого модуля и датчика. Более того, есть возможность дополнительно подключить светодиод, ЖК-дисплей или другие устройства вывода, чтобы выявить корректность работы модуля или датчика.

Тестирование

В качестве сравнительных тестов с аналогами подобных сборок проведены следующие тесты по основным параметрам:

- тест производительности процессора (ЦПУ);

- тест оперативного запоминающего устройства (ОЗУ);
- тест постоянного запоминающего устройства (ПЗУ).

Тестирование образа проходило в несколько итераций с целью подтверждения достоверности полученных результатов и отладки удобства использования для конечного пользователя.

В качестве инструмента для замера искоемых параметров для ЦПУ и ОЗУ использовался UnixBench Test. Это оригинальный набор тестов, который на протяжении многих лет обновлялся и пересматривался многими людьми. Основная цель этого инструмента – предоставить базовый индикатор производительности Unix-подобной операционной системы. UnixBench можно использовать для оценки производительности системы при выполнении одной или нескольких задач. Следует отметить, что это инструмент для тестирования системы, а не только инструмент для тестирования ЦП, ОЗУ или диска. Результаты будут зависеть не только от характеристик и параметров оборудования, но и от операционной системы, библиотек и даже компилятора [3].

В Unixbench просматривается большое преимущество предложенной сборки над конкурирующей. Наибольшая разница наблюдается в задаче shell scripting, где производительность увеличилась на 3571,91% выше, что означает, что теперь можно запускать и останавливать больше процессов за то же время. (рис. 3).

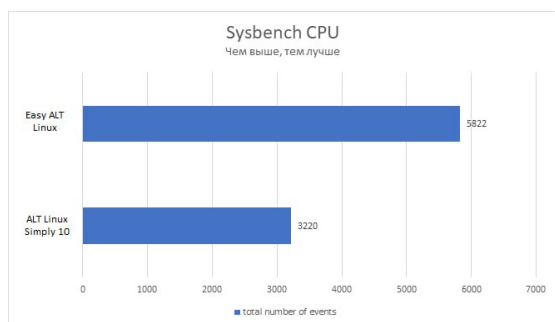


Рис. 3. Тест процессора

Аналогичное увеличение производительности во время теста ОЗУ на 123,17% больше в пользу новой сборки (рис. 4).

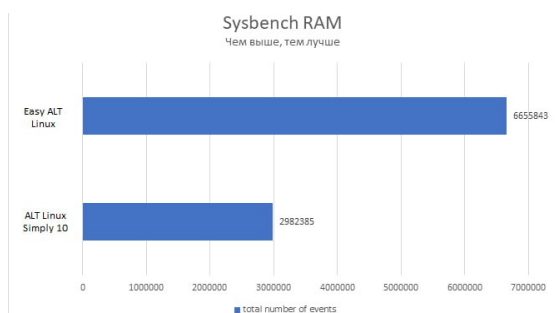


Рис. 4. Тест ОЗУ

Тест ПЗУ осуществлялся с помощью Bonnie Test. Bonnie – классический тест файловой системы и

устройств хранения. Данная утилита проверяет линейное посимвольное и блочное чтение и запись, а также шаблон перезаписи. Она также выполняет тест поиска. В то время как другие бенчмарки выполняют более сложные тесты, bonnie — очень надежная и портативная программа, подходящая для базового тестирования. При желании возможно протестировать операции с прямым вводом-выводом (O_DIRECT в Linux) [4]

Результаты испытаний ПЗУ показывают, что операции ввода\вывода, особенно в части интеллектуального чтения, также лидирует новый образ с отрывом показателя на 1173,08% (рис. 5).

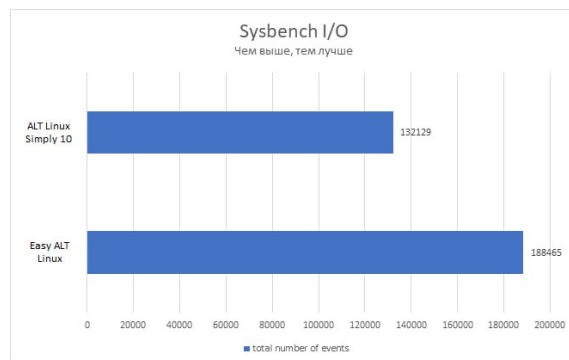


Рис. 5. Тест ПЗУ

Преимущества использование предложенного образа очевидны, производительность выросла в разы, а работа с ОЗУ и ПЗУ стала более оптимизирована.

Заключение

Данная работа имеет практическую значимость, создан новый дистрибутив ALT Linux для одноплатных компьютеров семейства Raspberry Pi 4.

В дополнение, в образ интегрирован авторский командный интерфейс, разработанный на языках программирования сценария Bash и высокоуровневого языка программирования Python, позволяющий упростить конфигурирование и настройку библиотек используемых при подключении датчиков и модулей [5,6].

Решение позволит инженерам использовать предложенный дистрибутив ALT Linux как базовый для создания и реализации более сложных задач и проектов, особенно в таких секторах, как производство, здравоохранение и транспорт.

Полученные результаты позволяют улучшить и добавить новые функции в текущую систему в рамках проекта «OpenSource», что будет полезно для целевого сообщества и имеет решающее значение для развития отечественного дистрибутива ALT Linux.

Проведенные экспериментальные исследования производительности, указывают на обоснованность выбранного решения, а также неоспоримый факт прироста производительности по таким основным параметрам как ЦПУ, ОЗУ и ПЗУ.

Список литературы

1. Управление пакетами ALT Linux [Электронный ресурс]. — Режим доступа:

https://www.altlinux.org/управление_пакетами#Установка (дата обращения: 09.04.2022).

2. ARM (архитектура) [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/ARM> (дата обращения: 16.04.2022).

3. Тест процессора (Unixbench Test) // GitHub [Электронный ресурс]. – Режим доступа: <https://github.com/kdlucas/byte-unixbench> (дата обращения: 19.05.2022).

4. Тест жесткого диска (Bonnie Test) // [bonnie++\(8\) - Linux man page](https://linux.die.net/man/8/bonnie++) [Электронный ресурс]. – Режим доступа: <https://linux.die.net/man/8/bonnie++> (дата обращения: 19.05.2022).

5. Создание скрипта на языке python для управления простыми функциями // GitHub [Электронный ресурс]. – Режим доступа: <https://github.com/avcourt/restful-pi> (дата обращения: 06.06.2022).

6. «Raspberry Pi Servo Motor control» // Tutorials for Raspberry PI [Электронный ресурс]. – Режим доступа: <https://tutorials-raspberrypi.com/raspberry-pi-servo-motor-control/> (дата обращения: 06.06.2022).

ПРОГРАММНЫЙ ЭМУЛЯТОР УСТРОЙСТВ ИНТЕРНЕТА ВЕЩЕЙ "VIRTUAL IOT LAB"

*Угодников А.А., Антипов М.М.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе рассматривается решение по разработке программного обеспечения, для запуска и работы с Arduino, Raspberry, STM32 в виртуальном пространстве, с возможностью запуска внешнего программного кода или алгоритмов.

Введение

В данный момент на рынке отсутствует эмулятор, позволяющий работать в виртуальном пространстве с микроконтроллерами Arduino, Raspberry, STM32, периферийным оборудованием и датчиками, а также позволяющий компилировать программный код непосредственно на платформе.

Условия карантина продемонстрировали, что подобный продукт может быть использован не только в учебном процессе, но также будет востребован при организации олимпиад, конкурсов хакатонов. Разработанный программный продукт может быть внедрен в образовательный процесс в НИУ ВШЭ для подготовки к практическим занятиям. Планируется внедрять данный программный продукт в рамках работы с абитуриентами и школьниками в рамках профориентационной деятельности, проектов Инженерный класс и IT-класс.

Постановка задачи

Перед нами стояла задача продолжить разработку программного обеспечения и внедрить различные функции, как визуальные, так и функциональные.

1. Создание подсистемы управления виртуальными устройствами: создание устройства, удаление устройства, запуск устройства на выполнение, назначение ресурсов и установка ограничений, создание сетевого взаимодействия для развертывания и взаимодействия с эмулятором Raspberry Pi.

2. Реализовать подсистему для преподавателя, которая позволит производить контроль и проверку разрабатываемого программного кода, а также посмотреть артефакты в проекте учащегося.

3. Реализовать подсистему для реализации пользовательских дополнений в виде датчиков, сенсорных систем, а также других элементов, способных подключаться к системе.

4. Расширение базы данных и библиотеки базовых компонентов: добавление новых микроконтроллеров и одноплатных компьютеров, датчиков и сенсоров.

Обзор существующих решений

На данный момент на рынке уже существует ряд сервисов, в которых пользователь может эмулировать работу микропроцессоров и различных сенсоров, но все они имеют различные недостатки и недоработки, по сравнению с предлагаемым нами решением.

Виртуальная лаборатория IoTIFY содержит набор виртуальных контроллеров и сенсоров, но там не получится поработать с несколькими виртуальными платами одновременно, платформа разрабатывается и не распространяется свободно. Также интерфейс данного сервиса является не самым удобным для пользователя, а запуск проекта занимает внушительное время.

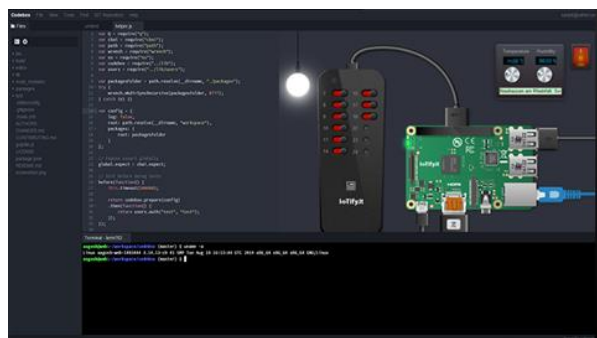


Рис. 1. IoTIFY Virtual Lab

Среда Tinkercad была включена в список существующих решений, так как позволяет работать с датчиками и Arduino, но Tinkercad не нацелен для создания IoT-проектов, в проекте можно использовать максимум одну плату, поддержка других микроконтроллеров и виртуальных плат не планируется.

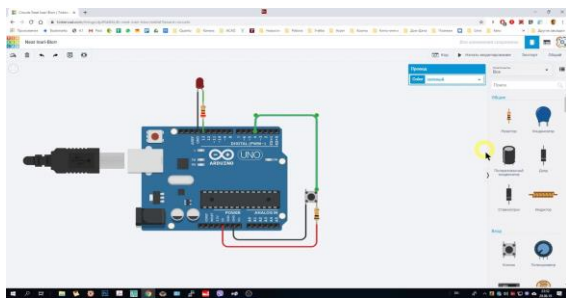


Рис. 2. Tinkercad

Также существуют другие решения, находящихся в открытом доступе, но все они также не могут позволить пользователю гибкую настройку проекта, возможность создавать собственные системы, а также качественно контролировать выполнение задания.

Технические параметры

Используемые технологии: Angular (JavaScript, HTML, CSS), Flask (Python), Docker.

Рабочий язык ПО: русский.

Вид ПО: веб-интерфейс с доступом к виртуальным платам и дополнительным компонентам, используя протокол SSH и выполняя POST-запросы на сервер.

Пользователи: ученики, студенты.

Особенности разработки ПО

Для разработки веб-интерфейса сайта используется фреймворк Angular, так как он заимствует некоторые лучшие аспекты разработки на стороне сервера и использует их для расширения возможностей разметки HTML в браузере [1]. В нем есть достаточно много полезных готовых решений, более продуманная система сбора и хранения информации, что является очень нужным при разработке данного проекта. Angular упрощает конструирование веб-сайтов и обеспечивает надежное их функционирование.

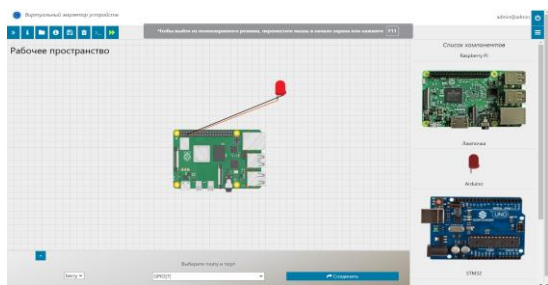


Рис. 3. Веб-интерфейс сайта для работы с виртуальными платами

Серверная часть использует фреймворк Flask на языке программирования Python. Используя Flask, можно выбирать, какие его компоненты будут применяться в приложении, и даже писать собственные [2]. Для управления базами данных в фреймворке Flask используется библиотека Flask-SQLAlchemy, которая позволяет выполнять различные операции с БД: создание таблиц, работа со строками. Созданная база данных сохраняет результаты работы пользователей, что позволяет реализовать систему для мониторинга

заданий, где преподаватель может посмотреть содержимое проекта, а также различные артефакты, такие как: исполняемые файлы, реализацию проекта и используемые микроконтроллеры и сенсоры.

Для выполнения обновления структуры базы данных существует библиотека Flask-Migrate. Схема базы данных, созданная для данного программного обеспечения, приведена на рисунке 4.

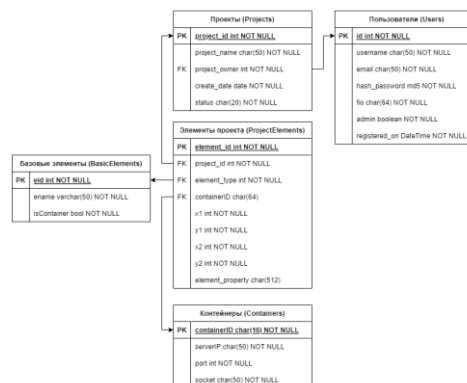


Рис. 4. Схема базы данных

Веб-служба для данного программного обеспечения будет в RESTful стиле. В данном проекте используется библиотека Docker SDK, которая позволяет работать с контейнерами, используя язык программирования Python [3]. На рисунке 5 изображена UML-диаграмма, которая описывает взаимодействие между пользователем и сервером.

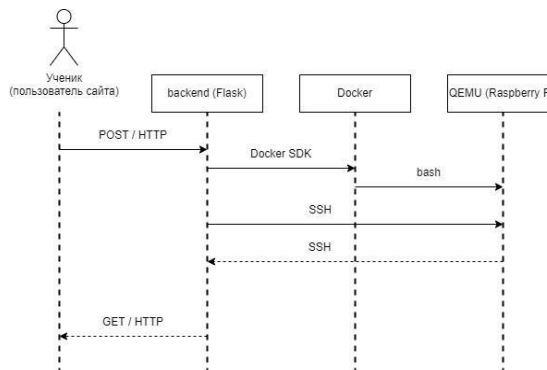


Рис. 5. UML-диаграмма взаимодействия пользователя с виртуальной платой

В рамках данного проекта была проведена исследовательская работа по анализу возможностей запуска Raspberry Pi, используя контейнеризацию и виртуальную машину QEMU, которая поддерживает работу с ARM архитектурой. Контейнеры содержат в себе образ Raspberry Pi 3b, который запускается на последних версиях эмулятора QEMU, так как поддержка USB-NET для работы интернета и взаимодействия по SSH доступна с версии QEMU 5.1 [4].

Для того, чтобы система в контейнере запускалась и настраивалась самостоятельно, используется пакет exect и модуль python3-pexect для скриптов на языке программирования Python соответственно. Так как запуск происходит через эмулятор QEMU, то

возможно реализовать выбор образа для Raspberry Pi при добавлении в проект. Raspberry Pi также содержит GPIO - порты ввода-вывода общего назначения [5]. Python-скрипт, находящийся в контейнере, считывает значения и направления портов, отправляет значения на серверную часть. На рисунке 6 приведена схема взаимодействия контейнера с back-end частью программного обеспечения.

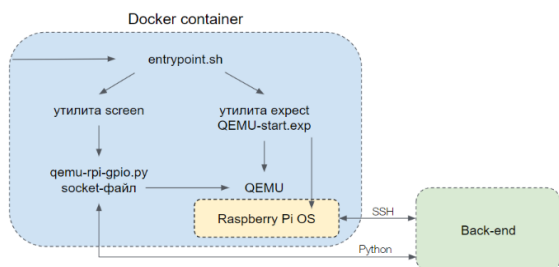


Рис. 6. Взаимодействие back-end части с контейнером Raspberry Pi

Для взаимодействия с Raspberry Pi через сайт мы используем SSH-протокол, тем самым мы позволяем пользователю подключаться к виртуальной плате, исполнять на ней команды, выполнять программы, которые загружены через сайт на виртуальную плату. Обмен файлами между Raspberry Pi и сервером реализован через утилиту scp.

Для взаимодействия между контейнером, backend и frontend частями проекта используются сокеты. Сокеты позволяют реализовать обмен между компонентами в реальном времени [6]. Так в контейнере, который работает на ОС Debian, выполняется скрипт на языке программирования Python. Данная программа выполняет ряд важных функций для создания и поддержания функционала контейнера: инициализация сокета для эмулятора QEMU, проверка изменения значений в соquete и функции для взаимодействия с backend-частью. Между frontend и backend частями также используются web-сокеты для передачи ввода из консоли от пользователя и отображения ответа сервера [7]. Таким образом пользователю не приходится дополнительно настраивать SSH соединение между своим устройством и контейнером, так как управление контейнерами полностью реализовано на backend-части.

Полную архитектуру программного обеспечения приведена на рисунке 7.

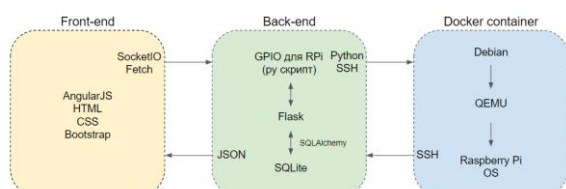


Рис. 7. Архитектура ПО

Заключение

Реализация данного программного обеспечения позволит пользователям создавать IoT-системы, используя виртуальные платы и компоненты, датчики,

изучать взаимодействие между одноплатными компьютерами и микроконтроллерами.

Список литературы

1. Фримен А. Angular для профессионалов. — "Издательский дом"" Питер""", 2017. — 800 с. — ISBN 9785446104512
2. Гринберг М. Разработка веб-приложений с использованием Flask на языке Python — М.: ДМК Пресс, 2016 — 274 с. — ISBN 978-5-97060-206-5
3. Docker SDK, URL: <https://docs.docker.com/engine/api/sdk/> (дата обращения 23.08.2022).
4. Raspberry Pi boards - QEMU documentation, URL: <https://www.qemu.org/docs/master/system/arm/raspi.html> (дата обращения 12.04.2022).
5. BCM2835 ARM Peripherals, URL: <https://www.raspberrypi.org/app/uploads/2012/02/BCM2835-ARM-Peripherals.pdf> (дата обращения 17.12.2021).
6. Python и DevOps: Ключ к автоматизации Linux. — (Серия «Бестселлеры O'Reilly»). Ной Гифт, Кеннеди Берман, Альфредо Деза, Григ Георгиу ISBN 978-5-4461-2929-4 - ЭБС Айбукс.ру [Электронный ресурс]. URL: <https://ibooks.ru/products/379920> (дата обращения: 13.12.2022).
7. Проектирование веб-API / Пер. с англ. Д. А. Беликова Лоре Арно ISBN 978-5-97060-861-6 - ЭБС Айбукс.ру [Электронный ресурс]. URL: <https://ibooks.ru/products/387363> (дата обращения: 26.06.2022).

АЛГОРИТМ МНОЖЕСТВЕННОГО ДОСТУПА С КОНТРОЛЕМ НЕСУЩЕЙ И ИЗБЕГАНИЕМ КОЛЛИЗИЙ (CSMA/CA) СЕТИ MESHTASTIC

Образцов О.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе оценивается пропускная способность сети беспроводных датчиков с использованием протокола Meshtastic. На основании марковской модели загруженности канала сравниваются алгоритмы множественного доступа к сети - линейный, реализованный в протоколе Meshtastic, и exponential backoff, описанный в стандарте IEEE 802.15.4. Цель работы - выявление недостатков реализации протокола Meshtastic и определение оптимального алгоритма множественного доступа к сети.

Введение

В последние годы большой интерес вызывает разработка беспроводных сетей для локальной связи. Базовая технология LoRa представляет собой протокол радиосвязи дальнего действия, доступный в

большинстве регионов и не требующий дополнительного лицензирования или сертификации. Поверх физического уровня LoRa разработан протокол самоорганизующейся (mesh) сети Meshtastic [2]. Узлы сети автоматически ретранслируют сообщения, которые они получают, чтобы создать ячеистую сеть для получения сообщения даже от самого дальнего устройства.

В данной работе оценивается пропускная способность сети Meshtastic в предположении идеальных условий канала для разных алгоритмов работы. При анализе мы предполагаем фиксированное количество станций, каждая из которых всегда имеет пакет, доступный для передачи. Другими словами, мы работаем в условиях насыщения, т.е. очередь передачи каждой станции всегда не пуста.

Анализ состоит в изучении поведения отдельной станции с помощью марковской модели [1] и получении стационарной вероятности того, что станция передает пакет в случайно выбранном интервале времени. Далее проводится сравнительный анализ алгоритмов exponential backoff [3] и линейного, представленного протоколом Meshtastic, в предположении модели Бьянки [1] с последующим выводом результатов.

Алгоритм множественного доступа к каналу в сети Meshtastic предполагает, что при передаче пакета с данными узел сети определяет, не ведет ли передачу одновременно другая станция, и в случае обнаружения такой коллизии узел откладывает передачу на случайный интервал времени в диапазоне $[0, W]$, где W - некоторый интервал времени (окно отсрочки). В Meshtastic данный интервал выбирается в зависимости от определенной узлом загрузки канала (доли времени, которую занимают передачи других станций) в соответствии со следующей формулой

$$W(p) = 2^{(1-p)CW_{min} + pCW_{max}}, \quad (1)$$

где значения параметров $CW_{min} = 2, CW_{max} = 8$.

По истечении времени ожидания попытка передачи повторяется.

В других стандартах радиосвязи, например, IEEE 802.11 и IEEE 802.15.4 применяется алгоритм с экспоненциальным увеличением окна отсрочки, при каждой неуспешной попытке передачи окно увеличивается в 2 раза, пока не достигнет некоторого максимального значения. Анализ алгоритма приведен в статье [5], приведем ниже основные положения данной модели.

Марковская модель

Рассмотрим фиксированное число n конкурирующих станций. Пусть каждая станция сразу же имеет пакет, доступный для передачи, после завершения каждой успешной передачи. Обозначим $b(t)$ за стохастический процесс, представляющий размер окна отсрочки для рассматриваемой станции в момент времени t .

Определим backoff window как - $W_i = 2^i W$, где $i \in (0, m)$ - стадия отсрочки, а $s(t)$ - стохастический

процесс, представляющий стадию отсрочки $(0, m)$ станции в момент времени t .

Ключевым приближением в модели Бьянки является то, что p - вероятность коллизии переданного пакета - не зависит от состояния $s(t)$ станции. В этом условии двумерный процесс $\{b(t), s(t)\}$ представляет собой цепь Маркова (рис. 1).

$$\begin{cases} P\{i, k | i, k + 1\} = 1, k \in (0, W_i - 2) \\ P\{0, k | i, 0\} = \frac{(1-p)}{W_0}, k \in (0, W_0 - 1) \\ P\{i, k | i - 1, 0\} = \frac{p}{W_i}, k \in (0, W_i - 1) \\ P\{m, k | m, 0\} = \frac{p}{W_m}, k \in (0, W_m - 1) \end{cases} \quad (2)$$

Вероятности перехода (2) учитывают, соответственно: 1) уменьшение счетчика времени отсрочки; 2) тот факт, что новый пакет после успешной передачи начинается с этапа отсрочки 0; 3) и 4) после неудачной передачи на i -м этапе отсрочки интервал отсрочки выбирается равномерно в диапазоне $(0, W_{min(i+1, m)})$.

Пусть $b_{i,k} = \lim_{t \rightarrow \infty} P(\{s(t) = i, b(t) = k\})$,

$i \in (0, m), k \in (0, W_i - 1)$ стационарное распределение цепи.

В силу цепных закономерностей имеют место следующие соотношения:

$$\begin{aligned} b_{i,0} &= p^i b_{0,0}, i \in (0, m-1), \\ b_{m,0} &= \frac{p^m}{1-p} b_{0,0}, \\ b_{i,k} &= \frac{W_i - k}{W_i} b_{0,0}, k \in (0, W_i - 1), \end{aligned} \quad (3)$$

Нормализация состояний через значение $b_{0,0}$ - $1 = \sum_{i=0}^m \sum_{k=0}^{W_i-1} b_{i,k} = \sum_{i=0}^m b_{i,0} \sum_{k=0}^{W_i-1} \frac{W_i - k}{W_i} = \sum_{i=0}^m b_{i,0} \frac{W_i + 1}{2} = \frac{b_{0,0}}{2}$

Пусть τ будет вероятностью того, что станция передает в обычном временном интервале. Поскольку любая передача происходит, когда окно отсрочки равно нулю, независимо от стадии отсрочки, это

$$\tau = \sum_{i=0}^m b_{i,0} = \frac{b_{0,0}}{1-p} = \frac{2(1-2p)}{(1-2p)(W+1)+pW}, \quad (4)$$

Заметим, что при выводе формулы (4) p фиксировано и в случае, когда $m = 0$, а W зависит от p , вероятность того, что станция начнет передачу в случайно выбранном слоте выражается следующим образом

$$\tau = \frac{2}{1+W}, \quad (5)$$

Заметим, что p - это вероятность того, что во временном интервале, по крайней мере, одна из $n - 1$ оставшихся станций передаст

$$p = 1 - (1 - \tau)^{n-1}, \quad (6)$$

где τ как функция от p (4). Численно решая (6), находим вероятность p , а следовательно, и τ . Как

только τ известно, вероятность P_{tr} того, что в интервале времени происходит по крайней мере одна передача, при заданных n активных станциях и вероятность P_s того, что передача успешна, легко получить как

$$P_{tr} = 1 - (1 - \tau)^n, \quad (7)$$

$$P_s = \frac{n\tau(1 - \tau)^{n-1}}{P_{tr}} = \frac{n\tau(1 - \tau)^{n-1}}{1 - (1 - \tau)^n}, \quad (8)$$

Произведение $P_{tr}P_s = n\tau(1 - \tau)^{n-1}$ есть доля временных слотов, в которые происходит успешная передача данных.

Определим пропускную способность системы – S , как долю времени, в течение которого канал используется для успешной передачи.

$$S = \frac{E[\text{время успешной передачи}]}{E[\text{время передачи}]} \quad (9)$$

Если $E[P]$ – средний размер передаваемых данных. Тогда $P_{tr}P_sE[P]$ – среднее количество информации, успешно переданной в единичный слот времени. Следовательно, пропускная способность системы $S \cong P_{tr}P_s$.

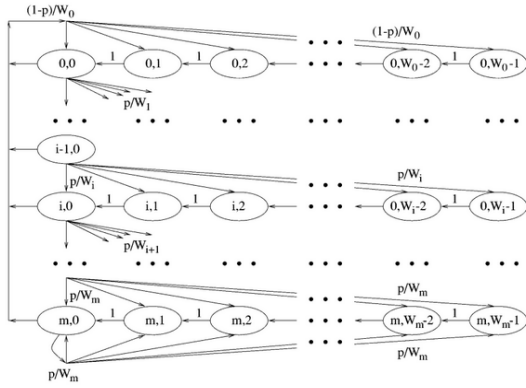


Рис. 1. Модель цепи Маркова для размера backoff window

Сравнительный анализ

1. Алгоритм CSMA/CA в Meshtastic

Предположим, что загруженность канала Meshtastic постоянна и равна p . В этом случае вероятность коллизии пакета в модели Бьянки также равна p . Размер окна W (backoff window) в Meshtastic зависит от p (из (1)).

Тогда согласно (6) и (1):

$$\tau = \frac{2}{1+W(p)} = \frac{2}{1+2(1-p)CW_{min}+pCW_{max}}, \quad (9)$$

Получаем следующую систему уравнений:

$$\begin{cases} \tau = \frac{2}{1+2(1-p)CW_{min}+pCW_{max}}, \\ \tau = 1 - (1 - p)^{\frac{1}{n-1}} \end{cases}, \quad (10)$$

2. Exponential backoff алгоритм CSMA/CA

В случае классической модели Бьянки для CSMA/CA алгоритма получаем:

$$\begin{cases} \tau = \frac{2}{1+W+pW^{\frac{1-(2p)^m}{1-2p}}}, \\ \tau = 1 - (1 - p)^{\frac{1}{n-1}} \end{cases}, \quad (11)$$

Первое уравнение системы имеет устранимый разрыв в точке $p = 0.5$, тогда

$$\begin{cases} \tau = \frac{2}{1+W+pW^{\frac{1-(2p)^m}{1-2p}}}, p \in [0, \frac{1}{2}) \cup (\frac{1}{2}, 1] \\ \tau = \frac{2}{1+W+pWm}, p = \frac{1}{2} \\ \tau = 1 - (1 - p)^{\frac{1}{n-1}} \end{cases}, \quad (12)$$

Результаты

Решая системы уравнений (10) и (12) относительно p при различных значениях n , построим графики зависимости загрузки канала и вероятности успешной передачи в зависимости от числа станций.

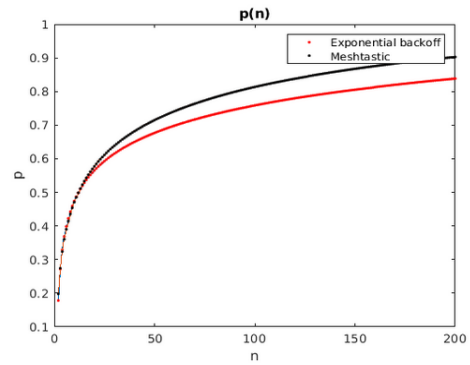


Рис. 2. График зависимости загрузки канала p от числа станций в сети n ($W = 8, m = 6, CW_{min} = 2, CW_{max} = 8$)

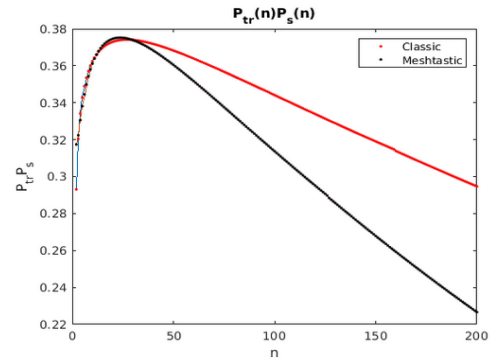


Рис. 3. График зависимости вероятности успешной передачи в случайный момент времени от числа станций в сети n ($W = 8, m = 6, CW_{min} = 2, CW_{max} = 8$)

Рис. 2 показывает, что при использовании алгоритма exponential backoff загрузка канала при том же количестве станций снижается по сравнению с оригинальным алгоритмом Meshtastic. При этом, как следует из рис. 3, вероятность успешной передачи данных при использовании алгоритма exponential backoff при том же количестве станций повышается, что позволяет сделать вывод о большей эффективности использования канала передачи данных.

Заключение

Таким образом, результаты анализа показали, что exponential backoff алгоритм в рассмотренном нами варианте (из 802.15.4), имеет большую пропускную способность по сравнению с алгоритмом протокола Meshtastic, в разумном интервале количестве устройств (от 2 до 100).

Заметим, что алгоритм экспоненциального увеличения окна отсрочки автоматически подстраивает размер окна в зависимости от загрузки канала. Варьируя параметр m , можно получить довольно широкие пределы, в которых оно может варьироваться, что позволяет алгоритму быть эффективным и при большом числе станций. Также алгоритм не требует постоянно включенного приемника, а значит, может быть использован в малопотребляющих устройствах, работающих на uplink.

Список литературы

1. G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function", IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 18, NO. 3, MARCH 2000.
2. Сайт проекта Meshtastic. [Электронный ресурс]. – Режим доступа: <https://meshtastic.org/> (Дата обращения: 25.01.2023 г.).
3. IEEE Computer Society, IEEE Standard for Local and metropolitan area networks - Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs). IEEE, 3 Park Avenue, New York, NY 10016-5997, USA, 5 September 2011.
4. T. To, A. Duda, "Simulation of LoRa in NS-3: Improving LoRa performance with CSMA", in Proc. IEEE Int. Conf. on Commun. ICC 2018, Kansas City, MO, USA, 2018 (DOI: 10.1109/ICC.2018.8422800).
5. C. Buratti, R. Verdone, "Performance Analysis of IEEE 802.15.4 Non Beacon-Enabled Mode", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 58, NO. 7, SEPTEMBER 2009.

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ НАЛИЧИЯ УСТАНОВОЧНЫХ ТРАСС ЗАДАННОЙ ДЛИНЫ ДЛЯ НЕИНИЦИАЛЬНЫХ АВТОМАТОВ

Евтушенко Л.Г.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Конечные автоматы часто используются, как формальная модель для описания поведения дискретных систем, и в ряде случаев установочные трассы позволяют упростить/оптимизировать процесс пассивного тестирования. В данной работе приводятся результаты компьютерных экспериментов проверки наличия установочных трасс заданной длины в конечном автомате.

Введение

При пассивном тестировании программных и аппаратных систем одним из важных критериев является количество необходимых проверок на каждом шаге работы системы [1, 2]. Если поведение тестируемой реализации представлено в виде конечного автомата, то по наблюдаемой входе-выходной последовательности (трассе) можно оценить текущее состояние системы и провести проверки только для требований, необходимых в этом состоянии, или только в критических состояниях. Наблюдаемая трасса называется установочной [2], если после нее достигается единственное состояние, независимо от начального состояния системы. Соответственно, в этом случае количество необходимых проверок может значительно сократиться. Поэтому интересным является вопрос, как часто в системах, поведение которых описано конечными системами переходов, существуют установочные трассы.

Известные результаты по построению установочных последовательностей [3] используют алгоритм построения таких последовательностей на основе дерева приемников и рассматривают только детерминированные инициальные автоматы. Алгоритм был модернизирован в [4] для возможности применения для недетерминированных автоматов. В работе [5] было предложено рассматривать вместо установочных последовательностей установочные трассы, включающие также выходные символы автомата; и в настоящей работе мы используем алгоритм из этой работы.

В настоящей работе представлены экспериментальные результаты по наличию установочных трасс заданной длины для конечных автоматов. Полученные результаты показывают, как в зависимости от начальных условий (мощности входного/выходного алфавитов, количества состояний, длины установочной трассы) изменяется процент наличия установочных трасс. На основании полученных результатов можно сделать выводы, какое оптимальное значение длины трассы стоит выбрать для последующего проведения пассивного тестирования заданного автомата в зависимости от необходимых условий оптимизации.

Постановка задачи

Для случайно сгенерированного набора полностью определенных, наблюдаемых, возможно недетерминированных неинициальных автоматов с заданной мощностью входного и выходного алфавитов i и o и числом состояний s , имеющего нормальное распределение, экспериментально установить процентное соотношение числа установочных трасс заданной длины к общему количеству трасс такой длины в группе автоматов с параметрами (s, i, o) .

Проанализировав полученные результаты, сделать выводы об изменении процента наличия установочных трасс, в зависимости от мощности входного/выходного алфавитов, числа состояний и длины установочной трассы.

1. Определения и постановка задачи

Под неинициальным конечным автоматом (FSM) S [3] будем понимать четверку (S, I, O, λ_S) , где S это конечное непустое множество состояний; I и O – конечные не пустые входные и выходные алфавиты, а $\lambda_S \subseteq S \times I \times O \times S$ – отношение переходов. Если четверка $(s, i, s', o) \in \lambda_S$, то автомат в состоянии s , приняв на вход воздействие i , выдаст выходной символ o и перейдет в следующее состояние s' . По определению, неинициальный конечный автомат может начинать функционировать в любом состоянии множества S . Обычно автомат рассматривается как трассовая модель, отображающая последовательности одного (входного) алфавита в (подмножества) последовательностей другого (выходного) алфавита. Автомат S называется *детерминированным* если для каждой пары $(s, i) \in S \times I$ существует не более одной пары (o, s') , такой, что $(s, i, s', o) \in \lambda_S$, в противном случае автомат S называется *недетерминированным*. Автомат S является *полностью определенным* если для каждой пары $(s, i) \in S \times I$ существует хотя бы одна пара (o, s') такая, что $(s, i, s', o) \in \lambda_S$, в противном случае автомат S называется *частичным*. Автомат называется *наблюдаемым*, если для каждой тройки (s, i, o) существует не более одного состояния s' такого, что $(s, i, s', o) \in \lambda_S$.

Входо-выходная последовательность γ называется *установочной* трассой для полностью определенного автомата S , если γ -преемник множества всех состояний является *синглтоном*, т.е. содержит одно состояние [5]. Согласно определению, известно, в какое состояние перейдет автомат после наблюдения установочной трассы, независимо от начального состояния, в котором начиналось функционирование автомата. [5]

2. Алгоритм построения установочных трасс

Для проведения экспериментального исследования используется алгоритм построения установочных трасс, представленный в работе [5].

Вход: полностью определенный, наблюдаемый, возможно недетерминированный автомат $\lambda_S \subseteq S \times I \times O \times S$ и $l \geq 1$

Выход: Сообщение: для автомата S не существует установочной трассы длины l ;

б) множество всех кратчайших установочных трасс длины не более l для S .

Шаг 1.

Строим усеченное дерево преемников для автомата S . Узлы дерева преемников маркируются как подмножества множества S , корень дерева маркируется как множество всех состояний из S , а ребра – парами входной/выходной символ.

Рассмотрим узел, помеченный как множество Z на уровне j , $l > j \geq 0$, и пару входной/выходной символ i/o . Если в множестве есть состояние s , из которого есть переход в состояние s' , помеченный парой i/o , то состояние s' включается в множество Q , помечающее узел на уровне $j + 1$, в который ведет переход из узла Z на j -м уровне. Узел, помеченный множеством Q на уровне k , $0 \leq k \leq l$, является *терминальным*, если мощность Q равна единице или $k = l$.

Шаг 2. Если в дереве-преемнике нет ни одного пути, терминальное состояние помечено множеством Q , мощность которого равна 1, то выдается сообщение “не существует ни одной установочной трассы длины $\leq l$ ”, в противном случае множество установочных трасс, которые маркируют пути дерева до терминальных узлов, помеченных синглтонами, есть множество всех кратчайших установочных трасс длины не более l .

Известно [1], что любое продолжение установочной трассы для полностью определенного наблюдаемого автомата также обладает этим свойством, если автомат является полностью определенным и наблюдаемым. Таким образом, на самом деле, приведенное выше дерево преемников представляет все установочные последовательности длиной не более l . Более того, имеет место следующее утверждение.

Утверждение 1. Для полностью определенного наблюдаемого, возможно недетерминированного автомата отношение числа установочных трасс длины $l \geq 1$ ко всем трассам такой длины не больше такого отношения для трасс длины $l + 1$.

Пример. Рассмотрим полностью определенный, наблюдаемый недетерминированный неинициальный автомат S на рис. 1.

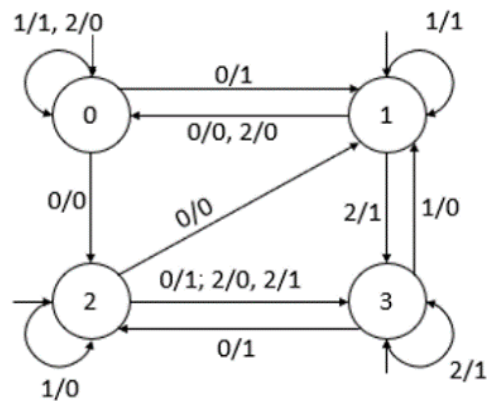


Рис. 1. Полностью определенный, наблюдаемый недетерминированный неинициальный автомат S

Построим дерево приемников для данного автомата. Фрагмент построенного дерева преемников представлен на рис. 2.

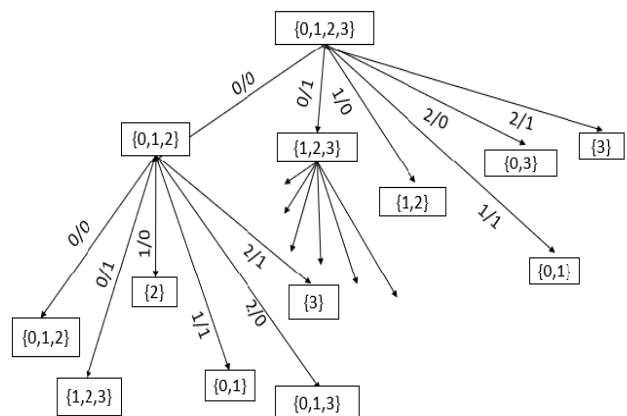


Рис. 2. Фрагмент дерева преемников

Количество всех трасс длины l равно $|I|^l * |O|^l$. При $l = 1$ число установочных трасс равно 1 (из 6), что составляет 16, 66 % от числа всех входо-выходных последовательностей длины 1.

При $l = 2$ число установочных трасс равно 21 из 36, что составляет 58,61 % от всех последовательностей длины 2. Согласно утверждению 1, при увеличении числа l процент установочных трасс может только увеличиться.

3. Описание экспериментов

Для проведения экспериментов, с помощью сервиса [6] генерируется набор из полностью определенных, наблюдаемых, возможно недетерминированных неинициальных автоматов с мощностью входного алфавита от 3 до 8, выходного алфавита от 3 до 8 и количеством состояний от 4 до 8, который имеет нормальное распределение. Алгоритм наследует алгоритм генерации детерминированных автоматов из работы [7], в которой, в частности, показывается, что для сгенерированного набора автоматов имеет место нормальное распределение. Для каждого автомата строится дерево приемников заданной высоты l и определяется процент существования установочных трасс длины l , где l изменяется от 1 до значения равного удвоенному числу состояний автомата. Вычисляется средний процент существования соответствующих установочных трасс для набора из 100 случайно сгенерированных автоматов.

Фрагмент результатов, отображающий процентное содержание установочных трасс с заданными параметрами на каждом уровне дерева приемников представлен на рис. 3.

i=3, o=3, s=3					
уровень	% установочных трасс				
1	37,5	33,3	55,5	50	55,5
2	75,47	69,4	83,8	81,8	79,6
3	90,03	85,6	94,1	93,5	91,1
4	95,80	93	97,9	97,7	96,2
5	98,22	96,5	99,2	99,2	98,4
6	99,23	98,2	99,7	99,7	99,3
i=3, o=3, s=4					
уровень	% установочных трасс				
1	22,2	11,1	0	55,5	33,3
2	63,7	60,5	58,4	77,9	58,8
3	84,61	83,3	81,7	88,2	78,8
4	93	92,8	91,7	91,5	88,7
5	96,7	96,6	96,2	93,5	94,1
6	98,4	98,4	98,2	96,5	96,9
7	99,2	98,7	99,1	98,9	98,3
8	99,3	98,8	99,2	99,4	98,9

Рис. 3. Фрагмент таблицы по процентному содержанию установочных трасс с заданными параметрами

Заключение

На основании проведенных экспериментов можно сделать вывод, что полностью определенные, возможно недетерминированные неинициальные автоматы обладают большим количеством установочных трасс, с помощью которых можно оптимизировать процесс пассивного тестирования.

Проанализировав, полученные в результате экспериментов данные, можно сделать выводы, что при длине установочной трассы, равной числу состояний автомата, в среднем 92% трасс являются установоч-

ными. При длине установочной трассы, равной удвоенному числу состояний автомата, процент установочных трасс приближается к 99%.

В качестве дальнейшей работы предполагается провести соответствующие эксперименты с автоматами, которые описывают поведение реальных компонентов телекоммуникационных систем, в частности, телекоммуникационных протоколов.

Список литературы

1. E. Bayse, A. Cavalli, M. Núñez, F. Zaïdi, "A passive testing approach based on invariants: application to the WAP", Computer Networks, Volume 48, Issue 2, 2005, pp. 247-266, <https://doi.org/10.1016/j.comnet.2004.09.009>.
2. N. Kushik, J. López, A. Cavalli and N.Yevtushenko, "Improving Protocol Passive Testing through "Gedanken" Experiments with Finite State Machines", 2016 IEEE International Conference on Software Quality, Reliability and Security (QRS), 2016, pp. 315-322, doi: 10.1109/QRS.2016.43.
3. Гилл А. Введение в теорию конечных автоматов // М.: Наука. – 1966. – 272 с.
4. Кушик Н.Г. Методы синтеза установочных и различающих экспериментов с недетерминированными автоматами: Диссертация на соискание ученой степени кандидата технических наук // Том. Гос. Ун-т. – 2013.
5. A. Tvardovskii, N. Yevtushenko., "Using homing traces for simplifying passive testing of discrete event systems", 2022 IEEE International Russian Automation Conference (RusAutoCon), 2022, doi: 10.1109/RusAutoCon54946.2022.9896240.
6. Test generation for Finite State Machine. URL: <http://fsmtestonline.ru/>
7. Дорофеева М.Ю. Исследование и разработка конечно-автоматных методов синтеза проверяющих тестов для управляющих систем: Диссертация на соискание ученой степени кандидата технических наук // Томский государственный университет. – 2007.

БИБЛИОТЕКА ДЛЯ СИНХРОНИЗАЦИИ РЕПЛИЦИРОВАННЫХ МИКРОСЕРВИСОВ

Минубаев М.И.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

Работа посвящена разработке библиотеки на языке python с открытым исходным кодом, предназначенной для синхронизации асинхронных реплицированных микросервисов.

Введение

В современном мире разработка программного обеспечения становится все более сложной и много-

аспектной. Рост числа пользователей и увеличение объемов данных требуют разработки масштабируемых и гибких систем. Одним из подходов к решению данной проблемы является использование распределенной микросервисной архитектуры. Микросервисы предлагают множество преимуществ, таких как модульность, гибкость и масштабируемость.

Ведущие авторы, такие как Martin Fowler, James Lewis [1] и Sam Newman [2], обозначают важность использования микросервисов для обеспечения эффективной работы распределенных систем. Кроме того, крупные компании, такие как Google, Netflix и Amazon, уже разрабатывают и используют свои решения для обеспечения синхронизации и согласованности данных в микросервисной архитектуре.

Python [3] - интерпретируемый высокоуровневый язык программирования, широко используемый в веб-приложениях, ориентированный на повышение производительности труда разработчика, мультиплатформенность и универсальность.

Асинхронное программирование [4] - парадигма, где системные вызовы становятся неблокирующими, позволяя продолжать работу программы без ожидания завершения операций ввода-вывода, требующих длительного времени выполнения.

Микросервисная архитектура [5] подразумевает под собой разделение монолитного сервиса на систему связанных между собой блоков (микросервисов), каждый из которых решает отдельную бизнес-задачу. Микросервисная архитектура позволяет масштабировать системы и увеличивать производительность без значительных трудозатрат.

Тем не менее увеличение количества реплик нарушает атомарность выполнения запросов, усложняет сообщение с внешними системами, вносит проблемы с консистентностью данных и нарушает конкурентность обработки запросов.

Существующие решения

Различные технологии и алгоритмы предлагают решения для обеспечения синхронизации и согласованности данных в распределенных системах. К таким технологиям относятся распределенные транзакции, такие как:

а) Two-phase commit (2PC) [6]: этот алгоритм используется для обеспечения атомарности транзакций в распределенных системах. Однако он может быть сложным в реализации, особенно в больших системах. Более того, он может стать узким местом производительности, так как требует двух раундов обмена сообщениями между всеми участниками перед выполнением транзакции.

б) Paxos [7] и Raft [8]: эти алгоритмы консенсуса используются для достижения согласованности в распределенных системах. Они обладают хорошими свойствами согласованности, но могут быть сложными в понимании и реализации. Они также могут столкнуться с проблемами производительности при большом количестве узлов или при наличии сетевых задержек.

в) Apache Zookeeper [9] и etcd [10]: это системы координации, которые предоставляют примитивы,

такие как блокировки и барьеры, для синхронизации операций в распределенных системах. Они обладают мощными функциями, но могут быть избыточными для некоторых приложений, которые требуют только простых механизмов синхронизации. Они также могут иметь ограничения по производительности при большом количестве узлов или при высокой нагрузке.

Многие компании активно используют микросервисную архитектуру для разработки своих систем. Однако, существующие решения для синхронизации и согласованности данных могут быть сложными, иметь ограничения по производительности и не всегда подходить для всех видов приложений и инфраструктур.

В связи с этим существует потребность в разработке новых подходов и инструментов, которые обеспечивают синхронизацию и согласованность данных между микросервисами и могут быть легко интегрированы в существующие системы и инфраструктуры. Данная статья описывает разработку Python-библиотеки, которая реализует три инструмента с использованием Redis для решения вышеописанных проблем: Распределенный Счётчик с Плавающим Временным Окном, Распределенный Уведомитель, Распределенный Ожидатель.

Технологический стек

Библиотека предназначена для языка python и разработана на нем, официальная поддержка заявлена для версий 3.9, 3.10, 3.11.

В разработке современных распределенных микросервисных архитектур важно выбрать подходящий инструмент для решения возникающих проблем с синхронизацией, консистентностью данных и обработкой запросов. В данной работе был выбран Redis в качестве основного инструмента, поскольку он обладает рядом преимуществ (масштабируемость, время отклика, отказоустойчивость), которые делают его подходящим решением для обработки этих задач.

Redis (Remote Dictionary Server) [11] является открытым, инструментом хранения данных в памяти с использованием структур данных ключ-значение. Redis обладает высокой производительностью и поддерживает разнообразные структуры данных, такие как строки, списки, множества, хеши, сортированные множества и геопространственные индексы.

Для взаимодействия с Redis используется асинхронная версия библиотеки redis-py.

Разработанные механизмы синхронизации

Данный раздел описывает как Redis используется для реализации механизмов синхронизации и их области применения в реальных системах. Это позволит понять, почему Redis является подходящим инструментом для решения проблем, возникающих в распределенных микросервисных архитектурах, и как он может быть успешно интегрирован в существующие и будущие проекты.

Распределенный счётчик с плавающим временным окном

Счетчик с плавающим временным окном позволяет отслеживать и контролировать количество со-

бытий, происходящих в определенном временном окне, и может быть использован для реализации ограничений скорости или механизмов регулирования нагрузки.

Распределенный Счётчик с Плавающим Временным Окном решает проблему контроля количества операций, выполняемых за определенный период времени, что особенно актуально для распределенных микросервисных архитектур, где важно контролировать нагрузку на отдельные сервисы и предотвращать их перегрузку.

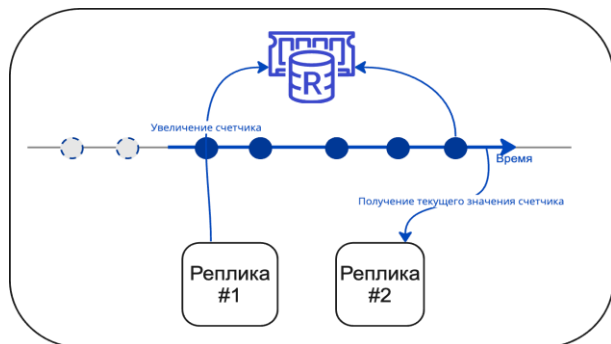


Рис. 1. Схема работы Распределенного Счетчика с Плавающим Временным Окном

В качестве примера использования распределенного счетчика с плавающим временным окном можно привести ограничение на количество запросов пользователя в секунду.

Для реализации механизма использовались сортированные множества (Redis Sorted Sets) – набор уникальных строк, упорядоченных по некоторому критерию [11]. Где строкой и критерием для сортировки является дата инкремента счетчика. При вставке элемента ему устанавливается срок жизни, равный временному окну, при получении текущего значения счетчика подсчитывается количество не истекших элементов, удовлетворяющих условию по нахождению в заданном временном окне.

Распределенный уведомитель

Распределенный Уведомитель — это механизм, обеспечивающий синхронизацию состояний между реплицированными микросервисами в реальном времени. Он гарантирует надежное и быстрое распространение уведомлений о событиях между репликами, что является критически важным для поддержания согласованности данных в микросервисной архитектуре.

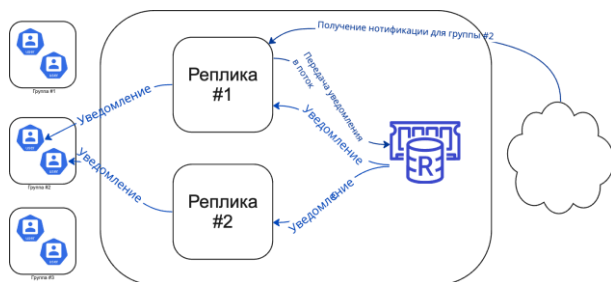


Рис. 2. Схема для Распределенного уведомителя

Пример применения: доставка уведомлений всем участникам группы, даже если они обслуживаются разными репликами. Это может быть полезно для оповещений о новых сообщениях, уведомлений об изменении статуса пользователей или информации о событиях.

Реализация основана на потоках (Redis Streams) – структура данных, разрешающая только вставку в конец списка [11]. Для каждого пользователя создается собственный поток. При получении нового сообщения, свойственного группе оно добавляется в каждый поток, связанный с этой группой, затем функции-слушатели получают сообщение из этого потока и используют по заданному сценарию.

Распределенный ожидатель

Распределенный Ожидатель — это асинхронная реализация механизма ожидания сигналов от других узлов или сервисов для организации обмена сообщениями между узлами. Он может подписываться на канал свойственный данной реплике и получать сигналы, а также уведомлять о произошедших сигналах другие узлы.

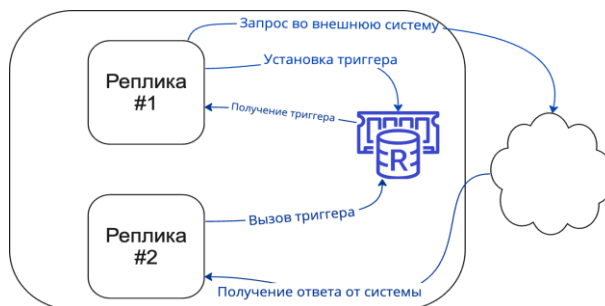


Рис. 3. Схема работы Распределенного ожидателя

Распределенный Ожидатель может быть использован для координации действий между репликами сервиса, позволяя им ожидать сигналы друг от друга перед выполнением определенных задач.

Реализация основана на парадигме издатель подписчик (Redis Pub/Sub) [11]. Распределенный ожидатель использует Pub/Sub для подписки на каналы, свойственные данной реплике, и получения уведомлений о сигналах от других узлов или сервисов.

Дальнейшее развитие

В дальнейшем планируется:

1. Расширить список механизмов синхронизации, предоставляемых библиотекой следующими классами:

а. Распределенные блокирующие задачи, гарантирующие выполнение за заданный период времени только на одной из реплик.

б. Распределенный искатель – механизм, позволяющий отправлять запрос с одной из реплик, а получать ответ на другой, включая случаи, когда ответ может быть получен ранее завершения запроса.

с. Распределенный кэш, позволяющий вычислять результат выполнения функции с заданными

аргументами единожды, а при последующих ее вызовах использовать ранее вычисленное значение

2. Провести нагрузочное тестирование каждого разработанного механизма, изучив возникающие задержки и их зависимость от интенсивности запросов.

Заключение

Redis является мощным инструментом для реализации механизмов синхронизации в распределенных системах. Он предоставляет эффективные и гибкие средства для ограничения скорости, контроля нагрузки, распространения уведомлений и организации обмена сообщениями между узлами. Использование Redis в микросервисных архитектурах позволяет достичь высокой производительности, надежности и согласованности данных

В рамках данной работы разработана python библиотека с открытым исходным кодом для синхронизации реплицированных микросервисов [12].

Библиотека значительно упрощает проектирование асинхронных масштабируемых систем, позволяя решать проблемы с репликацией данных, контролировать конкурентность при обработке запросов, избегать проблем с консистентностью данных.

Список литературы

1. Lewis J. и Fowler M., Microservices: A Definition of This New Architectural Term 2014. URL: <https://martinfowler.com/articles/microservices.html> (дата обращения 20.01.2023).
2. Newman S. Building Microservices / O'Reilly Media, 2015
3. Python documentation. URL: <https://docs.python.org/3/> (дата обращения: 20.01.2023)
4. Hatting C. Using Asyncio in Python: Understanding Python's Asynchronous Programming Features / C. Hatting. //California: O'Reilly Media, 2020.
5. Simon F. Python Microservices Development / F. Simon, Z. Tarek. // Birmingham: Packt, 2021.
6. Yousef A.-h. J. и Samaras G., Two-Phase Commit // Encyclopedia of Database Systems. Springer US. 2009. C.3204-3209.
7. Lamport L., Paxos Made Simple // Sigact News - SIGACT, T. 32, 2001.
8. Ongaro D. и Ousterhout J., In Search of an Understandable Consensus Algorithm // 2014 USENIX Annual Technical Conference (USENIX ATC 14), Philadelphia, PA, 2014.
9. P. a. K. M. a. J. F. P. a. R. B. Hunt, ZooKeeper: Wait-Free Coordination for Internet-Scale Systems // Proceedings of the 2010 USENIX Conference on USENIX Annual Technical Conference, Boston, MA, 2010.
10. L. a. T. W. a. K. C. a. E. E. a. K. M. Larsson, Impact of etcd Deployment on Kubernetes, Istio, and Application Performance // Softw Pract Expert, T. 50, C. 1986-2007, 2020.
11. Redis Inc., Introduction to Redis. URL: <https://redis.io/docs/about/> (дата обращения 20.01.2023).
12. Исходный код библиотеки. URL: <https://github.com/malik89303/aiodistributor/> (дата обращения 20.01.2023).

МЯГКОЕ ДЕКОДИРОВАНИЕ БЧХ КОДОВ

*Бандюков А.А., Метляков А.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе проводится анализ и оценка эффективности алгоритмов декодирования с мягким входом и мягким выходом для кодов БЧХ и расширенных кодов БЧХ для последующего использования данных методов декодирования в каскадных кодовых конструкциях. Результаты моделирования показывают, что данные методы имеют преимущества над классическими (hard decision decoding) способами декодирования.

Введение

БЧХ коды входят в классическую группу линейных кодов, которые хорошо себя показывают в исправлении ошибок [1]. По этой причине они широко используются в различных телекоммуникационных стандартах. Однако, в современных сетях все больше начинают использоваться каскадные кодовые конструкции, где применяются мягкие способы декодирования. Алгоритмы “распространения доверия” (BP) и в частности Sum Product (SP) демонстрируют высокую эффективность при декодировании кодов с малой плотностью проверок (LDPC) [2].

Для эффективного декодирования такими алгоритмами плотность проверочной матрицы должна быть как можно ниже, то есть количество ненулевых элементов должно быть минимальным, а также в ней не должно содержаться большое количество коротких циклов, так как в противном случае это приведет к распространению ошибок в BP. Проверочная матрица БЧХ кода не имеет такой структуры, так как у нее довольно большая плотность, поэтому к ним невозможно в явном виде применить методы мягкого декодирования, что препятствует их использованию в каскадных кодовых конструкциях с обменом мягкой информацией между кодами. По этой причине возникает необходимость реализации оптимального решения этой задачи относительно БЧХ кодов.

На данный момент существует несколько возможных способов для применения BP к блочным кодам. Для поставленной цели нами был выбран алгоритм низкой сложности преобразования проверочной матрицы БЧХ кода [3].

Цель: добиться возможности эффективного применения БЧХ кодов в каскадах, где используется мягкое декодирование.

Описание алгоритма

Данный алгоритм основывается на “уменьшении” и “разрежении” проверочной матрицы БЧХ кода. Изначально, проверочная матрица РСМ БЧХ кода имеет $n - k$ строк и n столбцов. Первый шаг рассматриваемого алгоритма - это расширение данной матрицы до квадратной формы ЕРСМ путем цикли-

ческого сдвига последней строки на один элемент вправо.

Следующий шаг алгоритма - это нахождение разреженного представления уменьшенной проверочной матрицы $RPCM$, то есть требуется преобразовать матрицу так, чтобы ее плотность уменьшилась. Это можно достичь путем линейного сложения 1-ой строки матрицы со всеми остальными, после чего на основе тех строк, что дают наибольшее уменьшение веса, выбирается новый вектор h . Данный вектор дает достоверное проверочное уравнение, так как оно было получено как линейная комбинация векторов исходной матрицы PCM . Полученный вектор h является базисом для построения матрицы $RPCM$.

$$H = \begin{pmatrix} h_k & \cdots & h_1 & h_0 & 0 & 0 & \cdots \\ 0 & h_k & \cdots & h_1 & h_0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & h_k & \cdots & h_1 & h_0 \end{pmatrix} \quad (1)$$

$$H_E = \begin{pmatrix} h_k & \cdots & h_1 & h_0 & 0 & 0 & \cdots \\ 0 & h_k & \cdots & h_1 & h_0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & h_k & \cdots & h_1 & h_0 \\ h_0 & 0 & \cdots & 0 & h_k & \cdots & h_1 \\ h_1 & h_0 & 0 & \cdots & 0 & h_k & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & h_1 & h_0 & 0 & \cdots & 0 & h_k \end{pmatrix} \quad (2)$$

Здесь (1) - исходная матрица БЧХ кода, (2) - матрица после циклического сдвига.

Последним шагом является создание разреженной матрицы $SPCM$. Для этого требуется взять -ое количество нулевых матриц, и случайным образом заполнить их элементами из матрицы $RPCM$ таким образом, чтобы в полученных матрицах не было столбцов, которые бы содержали исключительно нулевые элементы. Финальная матрица $SPCM$ получается путем конкатенации случайно заполненных n матриц.

Результаты моделирования

Был проведен анализ пяти матриц: $SPCM2$, $SPCM3$ и $SPCM4$ - это разреженные матрицы, разделенные на 2, 3 и 4 подматрицы соответственно; $RPCM$ - уменьшенная матрица после нахождения нового базиса с наименьшим весом; PCM - стандартная проверочная матрица БЧХ кода.

Для тестирования использовался классический алгоритм декодирования для LDPC кодов - "распространения доверия" в вариации Sum Product. Для сравнения результатов приведен график зависимости FER от SNR.

Для проверки корректности работы алгоритма была смоделирована передача сигнала по AWGN каналу в сочетании с модуляцией BPSK. Рассмотрим результаты тестирования алгоритма на примере кода БЧХ $N, K = 127, 71$.

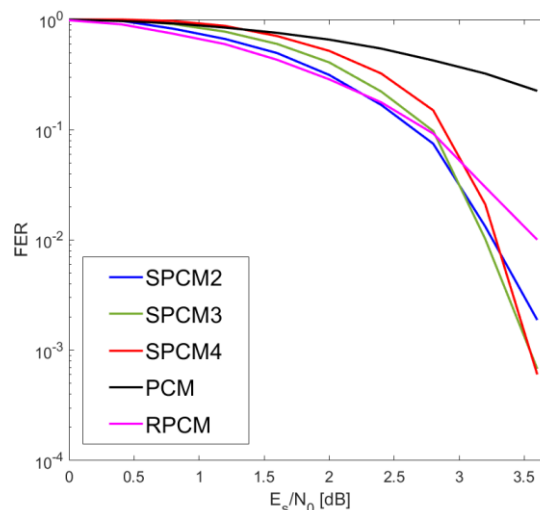


Рис. 1. Результат симуляции декодирования в канале (зависимость FER от SNR)

Из данного графика следует, что рассмотренный алгоритм мягкого декодирования дает значительно лучшие значения FER в зависимости от отношения сигнал шум для преобразованных матриц по сравнению с PCM . В случае с данным БЧХ кодом наилучший результат показывает $RPCM$ при низком отношении сигнал шум и $SPCM$, разделенная на 4 матрицы, при высоком.

Заключение

Таким образом, исследование показывает, что при поиске способов мягкого декодирования БЧХ кодов вполне возможно рассчитывать на успех.

Однако, стоит обратить внимание на то, что полученные результаты далеки от теоретически возможного идеала. В связи с этим, ближайшим продолжением работы будет решение задачи реализации алгоритмов предложенных в статье [4].

Список литературы

1. Сагалович Ю.Л., Введение в алгебраические коды, 2011.
2. Gallager, R. G., Low Density Parity Check Codes, Monograph, M.I.T. Press, 1963.
3. M. Baldi, F. Chiaraluce and G. Cancellieri, "Low-complexity soft-decision decoding of BCH and RS codes based on belief propagation.". URL: <https://www.researchgate.net/publication/236592169>
4. Kothiyal and O. Y. Takeshita, "A comparison of adaptive belief propagation and the best graph algorithm for the decoding of linear block codes.". URL: <https://ieeexplore.ieee.org/document/1523431>

СОЗДАНИЕ ОБРАЗОВАТЕЛЬНОЙ СРЕДЫ ДЛЯ ИЗУЧЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В РАМКАХ 2D ПЛАТФОРМЕРА

Котов А.В., Кадыргулов Т.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В данной работе представлен анализ поведенческих особенностей искусственного интеллекта (ИИ) в контексте 2D-платформера. Платформер используется как смоделированная среда, в которой агенты ИИ могут учиться и улучшать свою производительность методом проб и ошибок. Исследование сосредоточено на реализации алгоритмов обучения с подкреплением и влиянии различных вариантов дизайна на процесс обучения. Результаты демонстрируют потенциал платформера как эффективного образовательного инструмента для ознакомления учащихся с концепциями и методами машинного обучения с подкреплением. Полученные результаты способствуют разработке более привлекательных и доступных образовательных ресурсов для обучения искусственному интеллекту и машинному обучению.

Введение

Цель этого исследования — изучить потенциал платформера как инструмента обучения и разработать образовательную среду, которая включает алгоритмы и принципы машинного обучения в веселой и интерактивной форме. Научно доказано, что геймификация процесса обучения — эффективный способ изучения любой темы [1]. Платформер будет разработан для обучения основам искусственного интеллекта и машинного обучения, предоставляя учащимся теоретическое и практическое понимание предмета.

Актуальность исследования

Машинное обучение произвело революцию в том, как мы живем, работаем и общаемся в 21 веке [2]. От беспилотных автомобилей и персональных рекомендаций на платформах электронной коммерции до обнаружения мошенничества и диагностики заболеваний в сфере здравоохранения — алгоритмы машинного обучения стали неотъемлемой частью многих отраслей и приложений. Технология способна анализировать большие объемы данных и делать прогнозы, решения и рекомендации с уровнем точности, превосходящим возможности человека.

Поскольку технологии продолжают развиваться, спрос на специалистов с навыками машинного обучения быстро растет. Чтобы оставаться конкурентоспособными на современном рынке труда, становится все более важным хорошо разбираться в концепциях и методах машинного обучения. Это особенно актуально для людей, которые хотят сделать карьеру в области технологий, науки о данных и искусственного интеллекта.

Однако традиционные методы обучения искусственному интеллекту, такие как лекции и учебники, могут быть сухими и неинтересными для студентов. Создавая образовательные среды для изучения искусственного интеллекта в рамках 2D-платформера, преподаватели могут сделать изучение концепций искусственного интеллекта более доступным, интерактивным и увлекательным для учащихся.

2D-платформеры могут быть актуальными в данной теме, так как они обладают рядом преимуществ для образовательных целей. Во-первых, они легче в разработке и доступны для большего числа пользователей, чем более сложные 3D-игры. Во-вторых, они позволяют более простую и понятную визуализацию задач, что делает их более удобными для обучения искусственного интеллекта и машинного обучения. В-третьих, 2D-платформеры могут предоставлять разнообразные игровые ситуации, которые могут использоваться в образовательных целях, например, для обучения решению задач в области машинного зрения или навигации роботов.

Для анализа популярности запросов по теме исследования был использован инструмент Google Trends (см. Рис. 1 и Рис. 2).

Google Trends отображает соотношение числа запросов определенного слова или фразы к общему числу запросов в данном регионе за определенный промежуток времени. Затем эти данные нормализуются и приводятся к 100-балльной шкале. Наибольшее относительное значение получает 100 баллов, а остальные значения сравниваются с ним и принимаются на основе числа от 0 до 99.

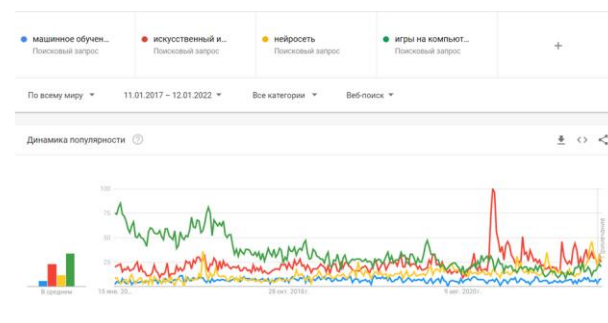


Рис. 1. Динамика популярности за последние 5 лет

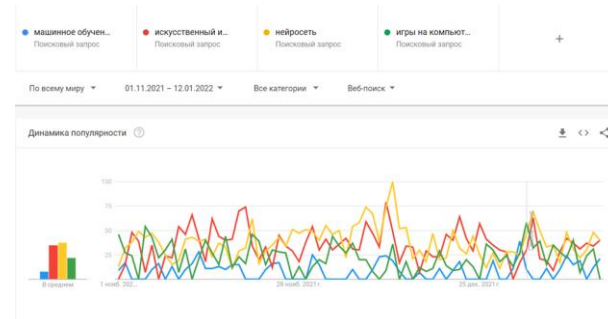


Рис. 2. Динамика популярности за последний год

Тема данной работы построена вокруг таких понятий, как «машинное обучение», «нейросеть», «игры на компьютер», а также «искусственный интеллект».

лект». Как мы можем заметить, количество запросов по данным темам растет. Средняя популярность запросов «нейросеть» и «искусственный интеллект» за последние 5 лет была 12 и 22, в то время как популярность запросов «игры на компьютер» составляла 34. Это показывает огромную популярность темы видеоигр. За последний год запросы «нейросеть» и «искусственный интеллект» поднялись до 38 и 35 соответственно, что в свою очередь показывает повышенный интерес к данной теме. Популярность запроса «игры на компьютер» упала до 22, что на общем фоне довольно высоко, но несмотря на это общая актуальность темы возросла, а актуальность данной ВКР и последующий рост интереса к ней.

Также стоит отметить популярность конкретно 2D-платформеров (см. Рис. 3). в рамках игровых жанров. Согласно динамике популярности за последние 5 лет, такие запросы как «2d game» и «pixel game» становятся все более популярными. Одной из причин высокого спроса является повышенное предложение на рынке игр: инди-разработчики игр стали еще более заинтересованы в создании двухмерных игр благодаря возможностям, которые предоставляют "народные платформы" по сбору средств, такие как Kickstarter. Разработка 2D игр оказывается гораздо более доступной и дешевой, что является определяющим фактором для маленьких, но энергичных команд разработчиков. К тому же, такой формат игры по-прежнему популярен у покупателей, что приводит к еще большей популярности 2D игр в индустрии.

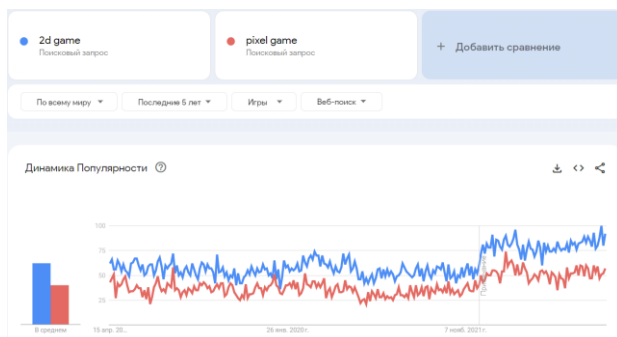


Рис. 3. Динамика популярности 2D игр за последние 5 лет

В заключение, данное исследование направлено на создание образовательной среды для изучения искусственного интеллекта с помощью 2D-платформера, что представляет большой потенциал для повышения мотивации студентов, удовлетворения спроса на знания в области искусственного интеллекта и является актуальным.

Аналоги

Для подтверждения важности и актуальности данной темы можно привести несколько примеров использования 2D платформеров в целях ознакомления с профессиональными сферами:

1. «Virtual patient» — игра-симулятор, которая позволяет врачам взаимодействовать с виртуальными пациентами и принимать решения на основе их

симптомов и истории болезни [3]. Это помогает врачам развивать навыки диагностики и решения проблем, а также улучшать свои навыки общения и ухода за пациентами. Эти типы игр не только развлекают, но и предоставляют медицинским работникам эффективный способ учиться и совершенствовать свои навыки в контролируемой среде с низким уровнем риска.

2. «Duolingo» - популярное приложение для изучения языков (см. Рис. 4), которое геймифицирует процесс обучения, используя различные игровые механики, такие как очки, награды и соревнования. Приложение начисляет баллы за правильные ответы на вопросы, и пользователи могут получать награды за прохождение уроков или достижение определенных этапов. Duolingo также включает в себя элементы соревнования, позволяя пользователям соревноваться с друзьями или другими пользователями Duolingo, чтобы увидеть, кто может прогрессировать быстрее или набрать больше очков. Кроме того, приложение использует игровые элементы, чтобы сделать процесс обучения более увлекательным и мотивирующим, такие как милые персонажи и забавная анимация. Используя эти игровые механики, Duolingo помогает пользователям оставаться мотивированными и вовлеченными в процесс изучения языка, облегчая им достижение своих целей.

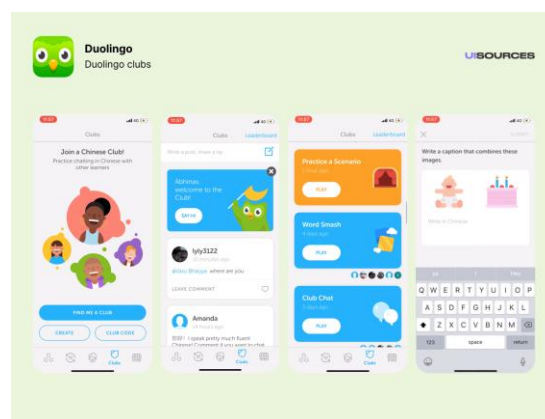


Рис. 4. Интерфейс приложения «Duolingo»

3. «Toca Lab» — это обучающая игра для детей, которая геймифицирует процесс изучения науки и техники. В игре игроки исследуют виртуальную лабораторию и используют различные инструменты для проведения экспериментов и изучения различных элементов и материалов. Игровой процесс разработан так, чтобы быть веселым и увлекательным, с красочной графикой, звуковыми эффектами и интерактивными элементами, которые побуждают игроков исследовать и открывать для себя. Играя в Toca Lab, дети могут развивать свое критическое мышление, навыки решения проблем и творческие способности, одновременно изучая научные концепции в игровой и увлекательной форме.

Это всего лишь несколько примеров многих образовательных сред для изучения определенных навыков в разных сферах на основе приложений, которые были созданы и используются для преподавания и

изучения необходимых концепций [4]. Сочетая элементы веселья, взаимодействия и вызова, эти образовательные среды обладают потенциалом для значительного повышения вовлеченности учащихся и их мотивации в изучении профессиональных навыков.

Предлагаемое решение

Разрабатываемое программное обеспечение позволяет облегчить процесс изучения ИИ, который часто может быть сложным и абстрактным процессом, особенно для тех, у кого нет опыта в области компьютерных наук или математики.

Приложение обладает следующими функциями:

1. Игра охватывает различные концепции машинного обучения, такие как обучение с учителем, обучение без учителя, обучение с подкреплением и глубокое обучение, а также показывает, как их можно применять в среде 2D-платформера.
2. Сервис предоставляет практический опыт работы с методами искусственного интеллекта с помощью испытаний и задач, что позволит игрокам применять то, что они узнали, в реальных сценариях.
3. Платформер предоставляет игроку визуальную обратную связь о производительности их моделей ИИ, помогая игрокам понять, как их ИИ принимает решения, и где его можно улучшить.
4. Наличие интерактивных учебных пособий, объясняющие концепции машинного обучения, алгоритмы и способы их реализации в игровой среде.
5. Приложение отслеживает прогресс игрока, позволяя ему видеть его рост и понимание с течением времени.

Включив эти функции, игра обеспечит комплексную образовательную среду для обучения машинному обучению, сочетающую элементы развлечения и игры с серьезным образовательным содержанием.

Реализация

Проект заключается в создании двухмерной игры-платформера, в которой механика игрового процесса основывается на принципах ИИ. В игре будет множество уровней, которые научат игроков различным аспектам ИИ, таким как машинное обучение и обработка естественного языка. Для закрепления знаний и понимания игра также будет содержать учебники и интерактивные викторины в образовательной среде.

Первым шагом был выбор игрового движка и создание оболочки в виде 2D платформера. На данный момент существует множество вариантов при помощи, которых можно реализовать поставленную задачу, например: Unity или Unreal Engine, однако движком была выбрана библиотека для языка python - Pygame [5]. Pygame — это популярная библиотека для разработки игр на Python, которая позволяет создавать 2D-игры. Pygame также имеет встроенную поддержку различных методов ИИ и машинного обучения, что делает его подходящей платформой для создания образовательной среды для изучения ИИ в 2D-платформере. Кроме того, Pygame предоставляет простой и доступный интерфейс для разра-

ботчиков, что делает его хорошим выбором для создания образовательного программного обеспечения.

Затем, после создания игры, следующим шагом будет интеграция ИИ-агентов в игру. Для этого можно использовать уже существующую структуру ИИ, например, Unity ML-Agents или OpenAI Gym, либо создать собственную систему ИИ.

Игровой процесс в данном случае фокусируется на пользователе, который играет за маленького человечка, настраивая его самостоятельно. В игре пользователь получает теоретические знания от учителей, которые предоставляют информацию из учебных пособий. После этого ему предлагаются практические задания для закрепления материала. Кроме того, в игре присутствует функция отслеживания прогресса, которая отображает выполнение пользователем заданий, предоставленных учителями, на разных уровнях игры. Таким образом, пользователь может отслеживать свой прогресс [6] в обучении концепциям ИИ.

Мы соберем данные об активности ИИ при помощи разных методов, включая записи игрового процесса, скриншоты и игровую телеметрию. Кроме того, мы будем анализировать действия и решения игроков, чтобы лучше понимать поведение и реакцию ИИ.

Для анализа данных мы будем использовать как количественные, так и качественные методы. Для количественного анализа мы будем применять статистические методы, чтобы выявить закономерности и тенденции в данных. Для качественного анализа мы тщательно проанализируем видеозаписи игрового процесса и телеметрию, чтобы понять поведение ИИ в различных контекстах.

Для получения дополнительных данных о влиянии поведения ИИ на игровой процесс мы будем собирать отзывы игроков с помощью опросов и интервью. Это поможет оценить восприятие игроками поведения ИИ и его влияние на игровой процесс.

После получения обратной связи от игроков, мы будем обеспечивать поддержание актуальности и работоспособности приложения, внедряя дополнительные функции и улучшения.

Заключение

Разработка образовательной среды для изучения искусственного интеллекта в рамках 2D-платформера является многообещающим подходом к изучению этой все более важной области. Используя увлекательный и интерактивный характер игр, учащиеся могут изучать концепции и методы машинного обучения в увлекательной и эффективной форме. Это может помочь улучшить мотивацию и удержание, а также сделать процесс обучения более доступным и приятным. Поскольку область искусственного интеллекта продолжает развиваться и расширяться, создание подобных образовательных сред будет играть все более важную роль в подготовке будущих поколений специалистов по искусственному интеллекту.

Список литературы

1. Zhan Z. et al. The effectiveness of gamification in programming education: Evidence from a meta-analysis //Computers and Education: Artificial Intelligence. – 2022. – С. 100096.
2. Wisetsri W. et al. An Investigation on the Impact of Digital Revolution and Machine Learning in Supply Chain Management //Materials Today: Proceedings. – 2022. – Т. 56. – С. 3207-3210.
3. Xu Y. et al. Learning experiences of game-based educational intervention in nursing students: A systematic mixed-studies review //Nurse Education Today. – 2021. – Т. 107. – С. 105139.
4. Hooshyar D. et al. An adaptive educational computer game: Effects on students' knowledge and learning attitude in computational thinking //Computers in Human Behavior. – 2021. – Т. 114. – С. 106575.
5. Politowski C. et al. Are game engines software frameworks? a three-perspective study //Journal of Systems and Software. – 2021. – Т. 171. – С. 110846.
6. Ferro L. S. The Game Element and Mechanic (GEM) framework: A structural approach for implementing game elements and mechanics into game experiences //Entertainment Computing. – 2021. – Т. 36. – С. 100375.

ПРОГРАММНЫЙ ПАКЕТ ДЛЯ ИСТОРИКОВ

*Конаныхин К.А., Поletaев Д.А.
Крымский федеральный университет
имени В.И. Вернадского,
Физико-технический институт,
кафедра радиоп физики и электроники*

Аннотация

Компьютерные технологии играют важную роль в научном и техническом документообороте, процессах моделирования и анализе данных. Однако применение информационных систем в сфере гуманитарных наук не ограничивается стандартным составлением текстовых файлов. Аспекты визуализации данных, подготовки презентаций также относятся к области интереса специалистов естественнонаучного профиля. В работе предлагается концепция пакета прикладных программ для историков. Описываются предлагаемые компоненты пакета, их взаимосвязь, аспекты работы и методология применения.

Введение

При интенсивном развитии компьютерной техники сложно назвать области науки, в которых не использовались бы информационные системы [1]. Их применяют физики и математики, архитекторы и филологи. Повышение быстродействия открыло широкие возможности для разработки дружественных к пользователю интерфейсов, осваиваемых непосредственно в процессе эксплуатации. Это существенно повышает удобство использования компьютерных систем и увеличивает широту охвата по отраслям науки. Так, специалистам в области филологии необ-

ходимы базы данных слов и их сочетаний, экономистам – адекватные модели, применяющие нейросетевые алгоритмы [2]. Исследователи же в области истории занимаются поиском и классификацией фактов, составлением карт, реконструкцией событий и описанием объектов. С одной стороны, им необходимо упорядочивать события и выстраивать эпизоды в хронологической последовательности, с другой – визуализировать картографическую информацию, проводить реконструкции сражений, внешних видов отдельных персонажей и изделий. Специалисты исторического факультета Московского государственного университета имени М.В. Ломоносова применяют, в основном, для данных целей кроссплатформенную среду разработки игр unity. Такой процесс оказывается весьма длительным, так как требуется отдельно разрабатывать примитивы, затем – применять их для визуализации. Кроме визуализации, накопленные сведения по исследованию целесообразно хранить в общем проекте с проведением индексации материалов для ускорения поиска и выделения соответствий.

Целью работы является формулирование концепции пакета прикладных программ для историков с описанием функциональных модулей и анализа их применимости.

Концепция

Современный пакет прикладных программ – это совокупность ряда приложений, связанных по принципу применимости к конкретной области наук [1]. Выделим основные требования, которые предъявляются историками к такому продукту: высокая скорость визуализации даже для неспециалиста в области информационных технологий, дружественный интерфейс, простота освоения, высокая скорость получения итогового результата, возможность обмена данными с другими приложениями. Рассмотрим каждый из этих пунктов по отдельности. Дружественный интерфейс предполагает простоту освоения программного продукта. Элементы управления требуются располагать так, чтобы начинающий пользователь осваивал базовые возможности за несколько сеансов работы. Для этого описываемый программный продукт требуется снабдить пошаговым руководством, в котором описываются основные приемы работы. Простота освоения предполагает минимальные усилия пользователя для старта выполнения большинства задач. Высокая скорость получения итогового результата подразумевает наличие гибко-настраиваемых модулей для различных задач. Возможность обмена с другими приложениями предусматривает доступ к экспорту и импорту данных из офисных и профессиональных программ. Структурная схема описываемого программного продукта представлена на рис. 1. На рисунке отмечены: блок ввода/вывода данных, модуль поиска совпадений, модуль визуализации, модуль картографирования, модуль импорта/экспорта данных, база данных, блок выдачи рекомендаций.

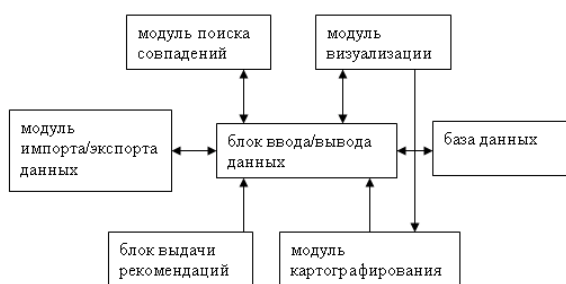


Рис. 1. Структурная схема пакета прикладных программ

Блок ввода/вывода данных осуществляет взаимодействие с пользователем посредством интерфейсов компьютерных систем. В качестве интерфейсов могут применяться компьютерная мышь, клавиатура, монитор, камера, принтер, сканер, 3d сканер, 3d принтер и др. Для упрощения процесса ввода данных предусматривается наличие наборов графических примитивов, объединенных в группы, а также – различных мастеров, производящих пошаговую подготовку данных, введение различных мастеров ввода данных, ряд поясняющих картинок и указателей. Блок ввода/вывода данных снабжается пошаговыми руководствами блока выдачи рекомендаций. Данный блок содержит иллюстрации по базовым операциям. Модуль поиска совпадений добавлен для поиска аналогичных элементов либо целостных структур в сети интернет, в других проектах. Поиск может осуществляться как по текстовым описаниям и ключевым словам, так и по графическим (трехмерным) моделям объектов. Для расширения диапазона поиска применяются нейросети [2] и различные статистические методы. База данных содержит описание проекта, ряд упорядоченных событий, данные карт и реконструкций сооружений, интерьеров, архитектурных ансамблей, сражений, а также ключевые слова для ускорения поиска того или иного материала. Модуль визуализации пакета прикладных программ содержит функционал для подготовки 3d структур из базовых блоков, легко-конфигурируемые элементы для отображения на картах (реконструкции сооружений, интерьеров, архитектурных ансамблей, сражений, движения войск). Кроме того, модуль визуализации связан с модулем поиска совпадений для обнаружения аналогичных элементов. Модуль картографирования предоставляет широкий функционал для подготовки карт, обозначения границ государств. Данный модуль предусматривает применение элементов, спроектированных в модуле визуализации. Модуль импорта/экспорта данных экспортирует данные в текстовые, графические форматы, а также производит считывание 3d объектов, разработанных в других приложениях.

Заключение

На текущий момент членами студенческого конструкторского бюро Физико-технического института Крымского федерального университета имени В.И. Вернадского предложена концепция графического интерфейса с наборами примитивов (простого для

освоения даже для неспециалистов в области информационных технологий), проанализированы возможности применения предлагаемого пакета, возможности его масштабирования, ведутся работы по разработке интерфейса.

Предлагаемый пакет прикладных программ может широко применяться как школьниками и студентами, так и ведущими исследователями.

Список литературы

1. Колодкин В.А. Основные этапы в истории возникновения и развития интернет-радио / В.А. Колодкин // Известия Южного федерального университета. Филологические науки. – 2014. – № 4.
2. Манжула В. Г. Нейронные сети Кохонена и нечеткие нейронные сети в интеллектуальном анализе данных / В.Г. Манжула, Д.С. Федяшов // Фундаментальные исследования. – 2011. – № 4.

РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДОВ ПОВЫШЕНИЯ ЭНЕРГЕТИЧЕСКОЙ ЭФФЕКТИВНОСТИ ГРАНИЧНЫХ УСТРОЙСТВ В СЕТИ IOT

Ильин А.Д.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии,
МИЭМ НИУ ВШЭ

Аннотация

Работа посвящена исследованию методов повышения энергетической эффективности граничных устройств в сети IoT. Анализируются подходы, выявляются их недостатки. Предлагается новый метод повышения энергоэффективности за счет использования специальных стратегий сбора и обработки данных для различных сценариев использования граничных устройств IoT.

Введение

Развитие устройств IoT и построение сенсорных сетей на их основе, позволяет открыть новые возможности в измерениях, мониторинге и управлении. Устройство IoT находят свое применение в стремительно расширяющемся списке областей: умный дом, транспорт, образование, научные исследования, умный город и т. д. Важную роль в их применении начинают играть такие методы как интеллектуальный анализ получаемых данных, интеллектуальное прогнозирование и сам сбор данных. Это позволяет обеспечить ранее недостижимые практические результаты.

Современные IoT устройства могут сильно отличаться друг от друга по типу связи, по объему и скорости генерации данных, энергозатратам, что существенно усложняет проектирование систем IoT. При увеличении требований к IoT системам они могут существенно усложняться и предъявлять повышенные требования к энергозатратам, что сводит на нет

их ключевое преимущество в простоте развёртывания и сроке работы от батареи.

Также важно отметить, что рост числа подключенных к сети устройств стремительно растет и по оценкам работы [1] может превысить планку в 30 миллиардов подключений. Это в свою очередь существенно увеличивает трафик данных, рост которого оценивают в 160 зеттабайт к 2030 году [2]. В результате для архитекторов IoT систем приоритетными становятся вопросы, связанные с эффективностью использования радиоканала, работы вычислительных систем, общей эффективности использования доступных энергоресурсов.

В настоящее время применяется архитектура централизованной обработки данных, которая уже сейчас сталкивается с недостаточной пропускной способностью радиоканалов, с уязвимостями в области безопасности и проблемами масштабируемости. Большие потоки данных, генерируемых сенсорами и датчиками IoT устройств, требуют развитие подходов, которые позволяют снизить объем передаваемых данных. В работе предлагается подход, основанный на проведении предварительной обработки данных ближе к самим устройствам сбора данных или на них самих. Они реализуются методами туманных и граничных вычислений. Однако использование этих подходов наталкиваются на ограниченные вычислительные ресурсы и энергопотребление граничных устройств.

Цель исследовательской работы заключается в исследовании и разработке методов, которые, позволят повысить энергетическую эффективность граничных устройств в сети IoT. Среди задач можно выделить: обзор предметной области, разработка и исследование предлагаемого метода повышения энергоэффективности граничных устройств IoT.

Методы повышения энергоэффективности IoT систем

Длительное время работы устройств IoT может быть достигнуто не только через оптимизацию и использование эффективных стратегий энергосбережения, но и с помощью различных методов сбора энергии от альтернативных источников питания [3]. Для примера можно взять солнечные панели, пьезоэлектрические или электромагнитные преобразователи, доказавшие свою эффективность для питания маломощных устройств IoT [4]. Отдельно можно упомянуть преобразование энергии радиоволн в электрическую.

Концепт беспроводной передачи энергии был известен почти столетие назад, однако несмотря на определённые достижения в технической сфере, данная технология все еще не смогла зарекомендовать себя как надежный и доступный альтернативный источник энергии. По этому направлению была представлена работа [5], где базовая станция IoT системы с некоторой периодичностью между отправками информационных пакетов посылает мощный сигнал, способный подзарядить оконечное устройство, снабженное устройством сбора электромагнитной энергии. В теории это позволяет вести сбор данных с устройств, лишенных батарейного питания. Это поз-

воляет радикально упростить обслуживание труднодоступных систем и повышает надежность. Однако данный метод существенно ограничивает использование радиоканалов, занимая потенциальное эфирное время.

При высокой загруженности радиоканала велик шанс получить коллизии при отправке данных. Оконечные устройства могут мешать друг другу отправлять данные при одновременном запуске передачи. Это приводит к потере данных, после чего необходима повторная отправка данных. Использование модуля радиосвязи на граничном устройстве является энергозатратным. В связи с этим необходимо сводить к минимуму его работу в активном режиме [6]. Для систем с малой дальностью связи эффективным методом борьбы с коллизиями является предварительное прослушивание канала передачи [7]. Недостатком данного подхода является низкая надежность при увеличении дальности в несколько километров за счет увеличения уровня шумов сравнимых с полезным сигналом.

Использование расписаний трансляций базовым модулем IoT сети, решает эту проблему, однако при этом существенно сужается область применения данных IoT систем, например, когда нужно иметь поток данных по запросу или в приложениях реального времени. В IoT системах используется широкий набор протоколов и технологий, например протокол LoRaWan более устойчив к коллизиям по сравнению с ZigBee, Sigfox, WiFi-ah, Z-Wave [7].

Следующий подход к повышению энергоэффективности граничных устройств IoT был упомянут в работе [8]. В подходе используется детальная проработка пути данных, начиная с самого сенсора и до отправки в облако или хаб с целью оптимизировать использование установленного микроконтроллера в устройстве IoT. Также проводится тщательная настройка времени и работы режима сна. Эти оптимизации являются основой энергоэффективной работы устройства. Однако при работе с задачами в реальном времени эффект значительно снижается. Более того в настоящее время наблюдается тенденция более активного использования вычислительных ресурсов оконечного устройства, так как их мощности и экономичность увеличиваются с каждым годом.

В работе [9] авторы обратили внимание, что стало целесообразным развёртывание облегченных библиотек машинного обучения на самих граничных устройствах. Это позволяет анализировать данные в реальном времени и соответственно отправлять на сервер готовое решение, а не массив данных.

Таким образом, на сегодняшний день было проведено много исследований по энергоэффективности и по поиску альтернативных источников питания, но если брать только один из вышеописанных аспектов в расчет, то вряд ли получится полностью решить проблему растущего числа оконечных устройств и их энергозатраты.

Метод повышения энергоэффективности

В этом исследовании предлагается новый метод повышения энергоэффективности оконечных

устройств за счет переноса вычислений на граничные устройства и адаптации методов и стратегий снижения энергозатрат под текущие потребности системы. Так в рамках задачи по мониторингу с фото-видео фиксацией можно использовать следующие стратегии сбора и обработки данных:

1. Фиксация факта движения на конечном устройством и отправка на облачную платформу соответствующего уведомления.

2. Запуск нейронной сети на конечном устройстве, распознавание объекта в кадре и отправка итогового результата на облачную платформу по команде.

3. Сбор на конечном устройстве статистики распознанных объектов и её периодическая отправка на облачную платформу.

Во всех случаях по каналу связи передаются пакеты малого размера, что позволяет использовать практически любую технологию передачи данных. Так ранее упомянутая LPWAN технология LoRaWAN, которая не была рассчитана передавать изображения, может применяться в данной системе. Что соответственно даст сэкономить энергию при использовании системы связи. Работа нейронной сети весьма энергозатратная по сравнению с простой переброской данных от сенсора до модуля связи, однако это может быть выгоднее, чем вести передачу данных в течение нескольких секунд, чтобы передать изображение. Так по приблизительным оценкам выигрыш может быть двухкратным.

Заключение

В исследовании была рассмотрена проблема энергоэффективности граничных устройств в сети IoT. Был проведен обзор предметной области и был предложен новый метод повышения энергоэффективности граничных устройств IoT, заключающийся в предварительной обработке данных на граничных устройствах IoT и использовании набора методов и стратегий снижения энергозатрат под каждый сценарий использования. Перенос обработки данных из облака на сами граничные устройства позволяет радикально снизить трафик используемых радиоканалов, а снижение объемов передаваемой информации, положительно сказывается на защите данных.

Список литературы

1. State of the IoT 2020 [Электронный ресурс]. URL: <https://iot-analytics.com/state-of-the-iot-2020-12-billion-iot-connections-surpassing-non-iot-for-the-first-time/> (дата обращения: 2.03.2022)
2. Alliance N. 5G white paper //Next generation mobile networks, white paper. – 2015. – Т. 1. – №. 2015.
3. Ju Q., Li H., Zhang Y. Power management for kinetic energy harvesting IoT //IEEE Sensors Journal. – 2018. – Т. 18. – №. 10. – С. 4336-4345.
4. Hu J. et al. Integrated data and energy communication network: A comprehensive survey //IEEE Communications Surveys & Tutorials. – 2018. – Т. 20. – №. 4. – С. 3169-3219.
5. Wang Y. et al. Energy-efficient data and energy integrated management strategy for iot devices based on

rf energy harvesting //IEEE Internet of Things Journal. – 2021. – Т. 8. – №. 17. – С. 13640-13651.

6. Rolich A., Alexander I., Voskov L. The study of Machine Learning Scenarios for the Internet of Arctic Things //2022 Moscow Workshop on Electronic and Networking Technologies (MWENT). – IEEE, 2022. – С. 1-7.

7. Pullmann J., Macko D. Increasing energy efficiency by minimizing collisions in long-range IoT networks //2019 42nd international conference on Telecommunications and Signal Processing (TSP). – IEEE, 2019. – С. 178-181.

8. Zhao S. et al. Understanding energy efficiency in iot app executions //2019 IEEE 39th international conference on distributed computing systems (ICDCS). – IEEE, 2019. – С. 742-755.

9. Varghese B. et al. Challenges and opportunities in edge computing //2016 IEEE international conference on smart cloud (SmartCloud). – IEEE, 2016. – С. 20-26.

ИССЛЕДОВАНИЕ И ОПИСАНИЕ РЕАЛИЗАЦИИ СИСТЕМЫ АВТОМАТИЧЕСКОЙ СЪЕМКИ МЕРОПРИЯТИЙ РАЗГОВОРНОГО ФОРМАТА

*Моисеев М.В., Галкина М.А., Мирталибов Д.И.,
Пивоваров А.С., Орлов К.О.*

*Национальный исследовательский университет
«Высшая школа экономики»,
департамента компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной статье раскрывается проблематика применения ручного трекинга при проведении съемок мероприятий разговорного формата. Помимо этого, в работе предлагается решение, реализованное в виде программы, позволяющей автоматизировать процесс слежения камерой за человеком при помощи системы компьютерного зрения.

Введение

В современном мире съемочный процесс является неотъемлемой частью создания видеоконтента и присутствует в большом количестве сфер жизни человека. Телевизионное производство, различные съемки в концертных залах и съемки в лекционных залах неуклонно набирают популярность на рынке [1], и одной из самых важных частей мероприятий в этих сферах является съемка спикеров, для которой наиболее часто встречаемым способом управления камерами является их ручной трекинг при помощи цифрового аналогового джойстика. Запись или трансляция как правило проводятся в течение нескольких часов, что является прямой причиной для роста частоты возникновения ошибок при съемках. На них в первую очередь влияет усталость оператора из-за длительной концентрации сил и внимания на рутинных действиях, а также существование «человеческого фактора», к которому можно отнести си-

туации, в которых у оператора возникают задержки с сопровождением спикера камерой или сменой кадра, а также ошибки и задержки в связи с наличием на локации недоступных для оператора ракурсов. Подобные случаи сопровождаются риском снижения точности съемки и уменьшения времени на дальнейший монтаж видео и решение творческих задач. Из этих пунктов вытекает вывод, что ручной трекинг на протяжении длительного времени напрямую влияет на снижение качества снимаемого контента.

Для решения этих проблем большинство организаторов съемок прибегают к увеличению количества сотрудников. Растет спрос на персонал, оборудование и прикладное программное обеспечение, которые не всегда способны решить основные проблемы. По этой причине существует необходимость автоматизировать процесс отслеживания человека в кадре с возможностью изменения цели и размера кадра путем создания системы референсного трекинга человека.

Предлагаемое решение представляет собой трекер, использующий машинное зрение и состоит из широкоугольной фиксированной камеры, которая направлена сразу на всю область съемки, и камеры класса видеоконференцсвязи (ВКС), которая будет управляться трекером. Камеры класса ВКС предлагается использовать по причине более точной механики по сравнению с более простыми камерами видеонаблюдения, а также из-за того, что они обладают возможностью использовать абсолютные координаты для ограничения рабочей области полем зрения широкоугольной статичной камеры и передачи им указания цели.

Трекер предполагает несколько способов управления, к которым относятся: управление с клавиатуры/пульта, управление при помощи приложения и голосовое управление. Такие способы наиболее удобны по причине привычного для режиссера повторения работы за пультом при наличии живых операторов на камерах с управлением ими по служебной связи.

Предложенное решение может быть использовано в сфере образования для записи лекций и семинаров, в корпоративной сфере во время записи собраний, совещаний и внутриорганизационных массовых мероприятий, а также в медиасфере для записи форумов, съездов, конференций и речей ведущих на концертах между выступлениями. Также оно может применяться для системы автоматической записи лекций для трекинга преподавателей [2].

Архитектура предлагаемого решения

Предлагаемое решение помимо используемых аппаратных ресурсов компьютера, на котором будет запущена программа, также требует наличия двух IP камер, способных отдавать RTSP-поток, а также поддерживающих протокол управления ONVIF. Такая конфигурация позволит работать с видеопотоками и с управлением камер удаленно через Интернет-сеть.

С программной точки зрения предлагаемое решение можно разделить на несколько компонентов.

Первый из рассматриваемых модулей отвечает исключительно за своевременное считывание видео с камер. Данный модуль должен позволять получать текущий кадр камеры в режиме реального времени (задержка в получении кадра не должна превышать 100 мс). После этого кадр со статичной камеры должен последовательно обрабатываться тремя компонентами: детектором, трекером и распознавателем.

Первый отвечает за обнаружение людей в кадре, после чего определяются их координаты относительно кадра. После этого, полученные объекты подаются на вход трекера, который назначает уникальный идентификационный номер каждому распознанному человеку. Благодаря этому появляется возможность отличать людей друг от друга на разных кадрах.

Наконец, распознаватель находит лица и сравнивает с локальным хранилищем известных программе лиц, позволяя определить личность человека (имя и роль в рамках съемки).

Используя все полученные данные, пользователь может выбрать цель съемки, используя либо идентификационный номер, либо имя человека в кадре. Это позволит отдавать команды на последний программный модуль - контроллер PTZ, который передает управляющие команды на движущуюся камеру по протоколу ONVIF.

Для работы детектора, трекера и распознавателя требуются различные нейронные сети: для нахождения тела человека, для присвоения идентификационных номеров, для нахождения и идентификации лиц. Использование таких моделей является достаточно ресурсоемким, в связи с чем было принято решение использовать клиент-серверную модель с тонким клиентом: все ресурсоемкие задачи будут выполняться на сервере, а клиент будет отвечать только за чтение потоков камер и за передачу команд управления.

Конечная схема архитектуры предлагаемого решения представлена на схеме (рисунок 1).

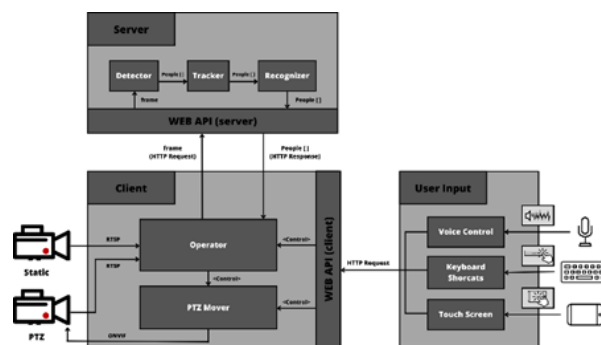


Рис. 1. Архитектура предлагаемого решения

Выбор моделей распознавания

Как было сказано ранее, в рамках предлагаемого решения должно быть использовано несколько нейронных сетей для обработки изображений. Каждая из них должна обладать высокой точностью, а

также наиболее высокой скоростью, иначе режим работы в реальном времени будет недостижим.

Первая используемая модель предназначена для обнаружения тел людей на полученном изображении. На данный момент наиболее распространенные и эффективные модели входят в семейство YOLO (You Only Look Once). Данные модели в своей базовой конфигурации обучены на датасете Microsoft COCO [3], включающем 80 классов, в том числе и класс "Человек". Наиболее продвинутые модели на данный момент относятся к группе YOLOv8 [4] (сравнение с другими группами приведено на рисунке 2), в связи с чем для предлагаемого решения стоит выбирать модель YOLOv8n - наиболее быструю и легковесную из группы.

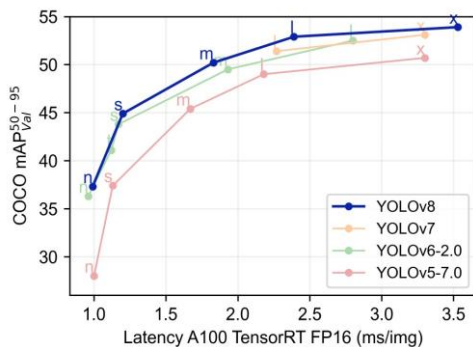


Рис. 2. Сравнение групп моделей, где скорость распознавания отмечена на горизонтальной оси, а точность на вертикальной

Другими ресурсоемкими задачами являются детекция и идентификация лица обнаруженного человека. Наиболее распространенное и эффективное на данный момент существующее решение - модуль face_recognition [5] для языка программирования python. Данный модуль может выполнять обе поставленные задачи, тем не менее в материале [6] рассказывается, что операция определения местоположения лица в данном модуле показывает низкие параметры точности и скорости. В связи с вышесказанным необходимо использование двух отдельных модулей - Facenet-pytorch [7] для детекции и face_recognition для идентификации. Такое сочетание позволяет достичь высокой скорости (в среднем на обработку фотографии двумя нейронными сетями затрачивается около 30 мс при использовании процессора Intel Core i5).

Алгоритм наведения камеры

Для корректного наведения камеры на человека используется комплексный подход, опирающийся на взаимодействие с дополнительной, референсной камерой. Для решения задачи наведения предлагается провести первичную настройку, а именно: навести PTZ камеру на левый и правый края сцены и запомнить значения угла поворота для дальнейшего использования.

После обнаружения цели в кадре статичной камеры, несложно вычислить относительную координату

положения человека разделив координату его середины на ширину кадра. Первой идеей для наведения камеры на человека может быть расчет угла поворота по формуле 1. Обозначим ширину кадра за x , координаты рамки человека $x1$ и $x2$, а предельные значения угла поворота за $\alpha1$ и $\alpha2$.

$$\alpha' = \frac{x2 - x1}{2x} * (\alpha2 - \alpha1)$$

Но, к сожалению, данная формула может привести к достаточно большой погрешности, которая будет увеличиваться по мере отдаления человека от центра кадра. Если же проводить более точные расчеты, то необходимо ввести также угол между сценой и камерой - β . Тогда угол поворота будет вычисляться по формуле 2. Тем не менее стоит понимать, что подсчет угла β неудобен для использования в программе и требует сложной настройки от пользователя.

$$\alpha' = \arcsin\left(\frac{\frac{x2 - x1}{2} - \sin\beta}{\sqrt{\left(x - \frac{x2 - x1}{2}\right)^2 + \frac{x^2 \sin^2 \beta}{\sin^2(\alpha2 - \alpha1)} - 2\left(x - \frac{x2 - x1}{2}\right) \frac{x \sin \beta}{\sin(\alpha2 - \alpha1)} (-\cos(\alpha + \beta))}}\right)$$

Взяв во внимание вышеописанную проблему, в предложенном решении рассматривается алгоритм, состоящий из нескольких шагов. Во-первых, камера будет наведена по формуле 1, с некоторой погрешностью. После этого человек будет определяться уже используя кадры с движущейся камеры, которая будет совершать небольшие движения в сторону человека, пока он не будет находиться в центре кадра (с определенной точностью).

Помимо горизонтального наведения требуется также подстраивать кадр и по высоте. Для решения этой задачи можно использовать схожий подход, за исключением шага с первоначальным наведением. Так, камеру можно двигать на некоторый шаг по высоте, пока линия глаз цели (которая определяется вместе с лицом) не будет с определенной точностью попадать на верхнюю треть кадра.

Наконец, предлагаемое решение также допускает возможность выбора требуемой крупности кадра. Выбор делается для одного из трех возможных значений: общий план, средний и крупный. Для получения желаемого эффекта используется тот же подход, что и в предыдущих двух случаях: приближение камеры сначала сбрасывается, а затем увеличивается с определенным шагом, пока высота рамки лица цели не попадет в нужный диапазон значений: для общего плана соотношение должно быть меньше или равно $\frac{1}{8}$, для среднего попадать в диапазон $[\frac{1}{4}, \frac{1}{3}]$, а для крупного $[\frac{1}{3}, \frac{1}{2}]$. Данные соотношения взяты в соответствии с пропорциями человека, описанными в [8].

Данный алгоритм не может выполняться для каждого кадра, так как в таком случае после удачного наведения все настройки будут сбрасываться и изменяться заново. Поэтому, предлагается проводить такую настройку единожды после выбора новой цели, после чего отслеживать исключительно измене-

ние координат человека, выполняя относительные движения в сторону смещения цели.

Тестирование

Для проверки гипотезы, основанной на возможности и преимуществе автоматизированного трекинга над ручным при проведении съемок и трансляций, было проведено тестирование разработанного модуля автоматического оператора. Данное тестирование проводилось на базе Медиацентра МИЭМ НИУ ВШЭ.

Для проведения эксперимента было сделано несколько тестовых записей: с использованием ручного трекинга и с применением разработанной системы для автоматизации трекера. При просмотре записей визуальных различий не было обнаружено. При проведении измерений было выявлено, что средняя задержка между моментом обнаружения человека и моментом наведения на него камеры составляет 45-50 мс, что доказывает возможность применения системы при проведении съемок в реальном времени. Данная задержка сравнима с той, что получается при проведении ручного трекинга. Но стоит учитывать, что время, потраченное на съемки тестовых записей, во много раз меньше времени проведения реальных съемок, при которых в силу усталости могут возникнуть ошибки со стороны оператора, а при применении автоматического трекинга показатели останутся неизменными и, как показало тестирование, эффективными.

Заключение

В результате проведенного анализа различных моделей для обработки изображений и распознавания людей как объектов была выбрана YOLOv8n как наиболее эффективная по скорости, весу и производительности. Для детекции лиц и их идентификации было принято решение использовать два отдельных модуля Facenet-pytorch и face_recognition, соответственно, для более высоких параметров точности и скорости распознавания.

Таким образом, в результате разработки и тестирования системы автоматической съемки мероприятий разговорного формата были получены данные, подтверждающие в целом возможность применения и преимущество автоматического трекера над ручным управлением с помощью джойстика. Данное решение позволяет сократить количество ошибок на съемках, возникших в результате «человеческого фактора», упростить и частично автоматизировать процесс съемок, в том числе увеличивая время на решение творческих задач и одновременно сокращая расходы на дополнительный персонал и оборудование.

Список литературы

1. Movie and Video Production - 2023 U.S. Market Research Report with Updated Recession Forecasts. Dec. 2022.
2. Cha Zhang, Yong Rui, Jim Crawford, Li-Wei He. An automated end-to-end lecture capture and broadcasting system. Jan. 2008.

3. Microsoft Common Object in Context (COCO). URL: <https://cocodataset.org/#home> (дата обращения: 23.01.2023).

4. Ultralytics YOLOv8. URL: <https://github.com/ultralytics/ultralytics> (дата обращения: 23.01.2023).

5. RetinaFace. URL: <https://github.com/serengil/retinaface> (дата обращения: 23.01.2023).

6. «Сравнение алгоритмов детекции лиц». URL: <https://habr.com/ru/post/661671/>. 18 Apr. 2022.

7. <https://github.com/timesler/facenet-pytorch> (дата обращения: 23.01.2023).

8. Harold B. Lee Library, Brigham Young University. Modelling; a guide for teachers and students. 28 June 2011.

RTSP-PROXY: ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ СУЩЕСТВУЮЩИХ РЕШЕНИЙ

Боярский А.О.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В IP-камерах для получения видеоданных используется RTSP протокол. Как показала практика, далеко не все камеры способны отдавать потоки в надлежащем качестве нескольким потребителям. Данная проблема решается с помощью RTSP прокси-сервера. В свободном доступе есть несколько приложений с функционалом RTSP прокси-сервера (далее RTSP-прокси), но нет материала, в котором оценивалась бы эффективность работы этих решений. Таким материалом и будет данная статья.

Введение

Real-Time Streaming Protocol[1] (далее RTSP) - прикладной протокол, предназначенный для получения аудио- и видеопотоков с IP-устройств. На качество выдаваемого по этому протоколу потока влияет множество факторов:

1. Настройки и качество IP-устройства, которое генерирует поток.

2. Настройки клиента – получателя данных по данному протоколу.

3. Устойчивость сети.

В Медиацентре МИЭМ на протоколе RTSP построено большая часть задач по съемке. Потребителями потока камер в сети МИЭМ может быть:

1. OBS, используемый для съемки и стриминга.

2. Система автоматической записи.

3. Система мониторинга.

4. NVR.

Практика использования IP-камер показала, что большинство из них не способно корректно работать с несколькими потребителями. Одни вообще не от-

дают видео больше чем одному потребителю, другие – рвут соединение со старыми потребителями при подключении новых. Данная проблема связана с малыми процессорными мощностями камер и решается переносом потребности работы с несколькими потребителями на RTSP-прокси[2][3]. В этой статье будет описан метод оценки эффективности работы свободно распространяемых решений, а также результаты эксперимента.

Постановка задачи

В открытом доступе есть 2 решения по проксированию протокола RTSP:

1. rtsp-simple-server[4]
2. live555-proxy-server[5]

Помимо них некоторые фреймворки позволяют выполнять задачи прокси-сервера. Это:

1. VLC
2. Gstreamer

Данные решения будут оценены по следующим метрикам:

1. Битрейту (должен быть приближен к 4096 кб/с, так как камеры, участвующие в эксперименте настроены на данное значение).
2. Задержке (должна быть меньше 100 мс, чтобы не была заметна человеческому глазу, что важно при стриминге онлайн).
3. Количеству камер, которые они могут одновременно проксировать.
4. Количеству клиентов, которым может отдавать поток.
5. Длительности непрерывного проксирования.
6. Способен ли проксировать все камеры (Медиацентр МИЭМ имеет 55 камер и кодеров в эксплуатации).
7. Гибкости (возможности работать по одному порту, выбору транспортного протокола, настройки задержки).

Помимо этого, есть вопрос оценки качества потока. Визуально ухудшение качества потока заметно по заторможенности видео и его пикселизации в разной степени. Это говорит о том, что природа данных эффектов может быть разная, поэтому инструмент оценки должен разделять каким-то образом причины между собой.

Ffmpeg как способ оценки качества работы прокси-сервера

Данный фреймворк можно использовать для оценки качества потока. Более того, так как и испытываемые приложения в большинстве своем могут использовать разные транспортные протоколы для потребления родного потока камеры, и сам Ffmpeg имеет данный функционал, в этом эксперименте будет также оценено влияние выбора транспортного протокола (TCP или UDP) для доставки RTSP-потока. Проксированный поток может быть записан с помощью команды `ffmpeg.exe -i --rtsp_transport tcp/udp rtsp://... -acodec copy -vcodec copy video.mkv`, а ошибки в получившееся записи можно получить с помощью команды `ffmpeg.exe -v verbose -i`

`.\video.mkv -f null ->error.log 2>&1`. Данная команда выводит три типа сообщений (ошибок):

1. Application provided invalid, non monotonically increasing dts to muxer in stream ...
2. Invalid POC ...
3. ... error while decoding ...

Первый тип сообщений свойственен записям, в которых некорректно выставлены отметки времени. Выражается это в прерывистости видеопотока. Связано это с тем, что при записи Ffmpeg непрерывно выставляет отметки времени. Каждая отметка соответствует своему кадру. Если связь с источником теряется на непродолжительное время (менее 7 секунд), при этом от источника не приходит сообщения EOS (End of stream), Ffmpeg продолжает запись, выставляет отметки времени, но не ставит им в соответствие кадры, из-за чего образуется эффект “дыры”. Визуально этот эффект можно наблюдать в плеере VLC. Например, если в десяти секундном ролике есть такая “дыра” длиной 2 секунды в середине, то плеер дойдя до 4 секунды записи, перескочит на 6, не проигрывая содержимое этих 2 секунд, что и будет образовывать данный эффект прерывистости. На графиках по результатам тестирования, такие сообщения будут под названием “DTS”.

Второй тип свойственен записям, в которых присутствует пикселизация или по-другому “сыплющаяся картинка”. Данный эффект можно наблюдать на рисунке 1.



Рис. 1. Визуальный эффект от потери RTP-пакетов

Возникает данный эффект по причине потери RTP-пакетов. На графиках по результатам тестирования, такие сообщения будут под названием “POC”.

Под третьим типом имеется ввиду группа сообщений, которые сигнализируют об ошибке декодирования. Визуальный эффект можно наблюдать на рисунке 2.



Рис. 2. Визуальный эффект от некорректной обработки входного потока прокси-сервером

Данный эффект возникает по причине некорректной работы прокси-сервера. На графиках по результатам тестирования, такие сообщения будут под названием “Error”.

Так как первые 2 типа ошибок имеют разную причину, но означают потерю RTP-пакетов в той или иной форме, то сумма количества данных ошибок будет показывать общий уровень стабильности соединения и объем потерянных данных. Это важный показатель, который влияет на качество производимого контента с использованием RTSP-прокси, значит наименьшее количество ошибок в совокупности с другими показателями и будет влиять на выбор наилучшего по производительности RTSP прокси-сервера.

Результаты исследования

Результаты исследования в Таблице 1.

Таблица 1. Результаты исследования

	Live555-proxy-server	VLC	Rtsp-simple-server	Gstreamer
Битрейт	4096 кб/с	4096 кб/с	4096 кб/с	4096 кб/с
Задержка	<100 м/с	>100 м/с	<100 м/с	<100 м/с
Настраиваемость задержки	Нет	Нет	Нет	Да
Количество источников	>1	1	>1	1
Количество потребителей	1	1	6	8
Длительность непрерывного проксирования	<5 минут	<5 минут	<16 часов	от 50 минут
Количество успешно проксируемых устройств (из 55)	19	19	52	55
Выбор транспортного протокола	есть	Нет	Да	Да
Выбор порта	Да	Да	Да	Да
Назначение пути	Нет	Да	Да	Да

Далее дана информация, поясняющая тонкости запуска/работы прокси-серверов, а также результаты эксперимента по оценке влияния транспортного протокола на качество проксированного потока.

Live555-proxy-server

Есть возможность выбора транспортного протокола и порта, на котором будет сервер будет принимать запросы.

Специфика эксперимента: так как количество входящих потоков влияет на качество исходящих (то есть проксированных), тестирование проводилось в режиме один запущенный экземпляр – один входящий поток для лучшего качества исходящего потока.

Результаты тестирования TCP-UDP на рисунке 3.

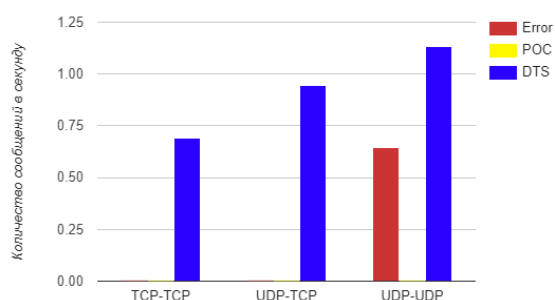


Рис. 3. Результат тестирования Live555-proxy-server

VLC

Тестирование проводилось с помощью консольной команды `cvlc -vvv <in> --rtsp-frame-buffer-size=500000 --sout=#rtsp{sdp=rtsp://:<out_port>/<out_path>} --sout-keep --sout-all --loop`. Так как для выполнения задачи проксирования VLC использует инструменты, созданные в рамках программного обеспечения Live555, с помощью которых и проксирует Live-555-proxy-server, результаты, ожидаемо, оказались схожими, их можно наблюдать на Рисунке 3. Результаты исследования:

1. Так как запускается через консоль, нужно запускать по экземпляру на поток.

2. Результаты тестирования UDP-UDP (так как работает только по UDP) повторяют результаты такого же тестирования Live555-proxy-server.

Помимо этого, есть свои особенности:

1. VLC имеет внутреннюю блокировку на запуск с правами суперпользователя, что не позволяет разворачивать инфраструктуру проксирования нескольких камер с помощью systemd.

2. В зависимости от версии параметр `--rtsp-frame-buffer-size` может как присутствовать, так и отсутствовать.

Rtsp-simple-server

Есть возможность исключения транспортных протоколов из использования, возможность назначения размера буфера (не влияющая на качество проксирования RTSP), назначении пути в ссылке прокси-потока, выбора порта.

Специфика эксперимента: так как количество входящих потоков влияет на качество исходящих (то есть проксированных), тестирование проводилось в режиме один запущенный экземпляр – один входящий поток для лучшего качества исходящего потока.

Результаты тестирования TCP-UDP на рисунке 4.

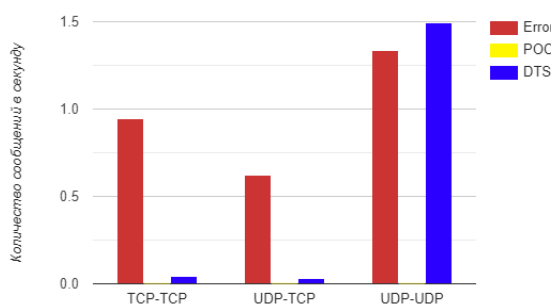


Рис. 4. Результат тестирования rtsp-simple-server

rtsp-simple-server в docker

Разработчик также предоставил возможность запуска своего прокси-сервера в Docker, собрав Docker-образ. Результаты тестирования TCP-UDP на рисунке 5.

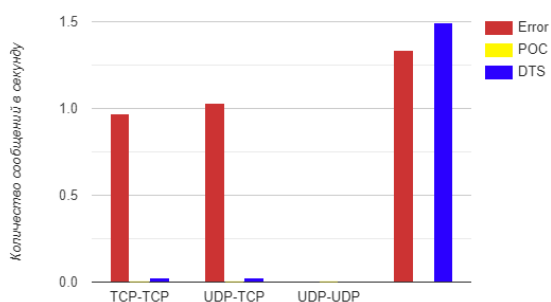


Рис. 5. Результат тестирования Rtsp-simple-server в docker

Gstreamer

Была собрана консольная команда `gst-launch-1.0 rtspsrc latency=50 location=<in> name=rtsp ! rtph264depay ! h264parse ! rtspclientsink latency=50 location=rtsp://<host_ip>:<port>/<path> name=rez rtsp. ! decodebin ! audioconvert ! voaacenc ! rez..` Особенности использования:

1. Плагин `rtspclientsink` не умеет открывать себе порт для прослушивания. При этом `Rtsp-simple-server` может это делать за данный плагин (`Rtsp-simple-server` даже распознает путь, указанный в `location`), поэтому экземпляры с данной командой можно запускать только на тех портах, которые слушают экземпляры `Rtsp-simple-server`.
2. Запускается через консоль, поэтому один входной поток на запущенный экземпляр.
3. До 8 потребителей на экземпляр (с пятого потребителя значительно увеличивается длительность соединения, до двух минут).
4. Стабильность соединения – от 50 минут до более суток (зависит от устройства).

Результаты тестирования TCP-UDP на рисунке 6.

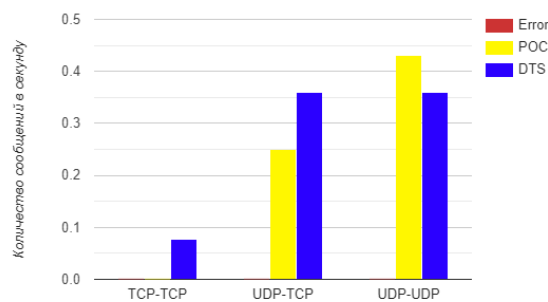


Рис. 6. Результат тестирования gstreamer

Выводы

Если говорить о численных значениях, то:

1. Ни один из прокси-серверов не меняет битрейт при проксировании.
2. У большинства серверов задержка удовлетворяет условиям.
3. `Gstreamer` является самым стабильным прокси-сервером по значению средней потери пакетов. При стабильной сети он теряет не более одного пакета в 10 секунд.

Помимо этого, результаты практического использования показывают:

1. Собственные инструменты проксирования в `Live555-proxy-server` и `Rtsp-simple-server` имеют свойство пренебрегать качеством обработки в угоду скорости обработки.
2. `VLC` не улучшает работоспособность инструментов проксирования `Live555`.
3. Дополнительный функционал `Rtsp-simple-server` позволяет использовать `Gstreamer`, как инструмент проксирования. Такой симбиоз выполняет функции прокси-сервера лучше всех выше представленных приложений, поэтому и была развернута в Медиацентре МИЭМ для постоянного использования.
4. Исследование эффекта от выбора транспортного протокола позволило опровергнуть одно из заблуждений в видеопроизводстве. При работе с протоколом RTSP часто не рекомендуют использовать TCP, но результаты тестирования показали обратное, что также было затронуто и в другом исследовании [6], где RTSP сравнивался с другими стриминговыми протоколами.

Список литературы

1. H. Schulzrinne, Columbia U., A. Rao, Real Time Streaming Protocol (RTSP). - 1998. - No. rfc2326.
2. Z. Shibeshi, A. Terzoli, K. Bradshaw Using an RTSP proxy to implement the IPTV Media Function via a streaming server // International Congress on Ultra Modern Telecommunications and Control Systems. - 2010.
3. Hui Guo, Guobin Shen, Zhiguang Wang, Shipeng Li Optimized streaming media proxy and its applications // Journal of Network and Computer Applications. - 2007. - Vol. 3-. - No. 1. - pp.265-281.

4. rtsp-simple-server | github : [сайт] - 2023. - URL: <https://github.com/aler9/rtsp-simple-server> (дата обращения: 15.01.2023). - Текст: электронный.

5. live555-proxy-server | live555 : [сайт] - 2023. - URL: <http://www.live555.com/proxyServer/> (дата обращения: 15.01.2023). - Текст: электронный.

6. A. Aloman, A.I. Ispas, P. Ciotirnae, R. Sanchez-Iborra, M.D. Cano Performance Evaluation of Video Streaming Using MPEG DASH, RTSP, and RTMP in Mobile Networks // 2015 8th IFIP Wireless and Mobile Networking Conference (WMNC). - 2015.

СРЕДА МОДЕЛИРОВАНИЯ АТАКУЮЩИХ СЦЕНАРИЕВ АЛГОРИТМА КОНСЕНСУСА TANGLE

Андреев А.Г.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент программной инженерии,
ФКН НИУ ВШЭ*

Аннотация

Программная составляющая данной работы представляет воспроизведение поведенческих особенностей системы распределённого реестра на основе Tangle относительно состояния агентов сети. Реализация заключается в автономности поведения агентов с обыденными и корыстными целями в окружении повседневной транзакционной рутины.

Основной ценностью самой работы являются описания вредоносных сценариев атак двойной траты в рамках алгоритма консенсуса Tangle. Реализация среды моделирования позволяет проводить эксперименты и сравнивать различные конфигурации транзакционной сети на предмет эффективности и устойчивости к поведенческим особенностям вредоносных и ведомых агентов.

Введение

В эпоху развития Internet of Things технологии систем распределённого реестра, гарантирующие быстроту, безопасность и распределённое распространение транзакций, являются особо востребованными. Одним из наиболее перспективных решений в мире криптовалют является МІОТА [1], которое основано на структуре данных в виде направленного ациклического графа с алгоритмом консенсуса Tangle [2]. Данное решение гарантирует линейную масштабируемость и безопасность сети при условии постоянной и высокой активности агентов, что не указывает на его универсальность и не даёт прямого представления о его возможностях применения.

Основа работы лежит в выявлении слабых сторон систем распределённого реестра на основе направленного ациклического графа ввиду нестандартных и непредвиденных поведенческих ситуаций. Программная составляющая работы предлагает различные потенциальные варианты решения этих про-

блем, которые опираются на исходные значения, поставленной перед системой задачи.

Результат данной работы заключается в описании ряда экспериментов, описывающих возможное усиление системы, над параметризированной транзакционной сетью на предмет ее безопасности относительно вредоносных агентов и потенциал ее применения.

Среда моделирования

Как уже говорилось выше, программная составляющая работы выражена в воспроизведении поведенческих особенностей различных агентов сети. Достигается это с помощью синхронной среды моделирования. Среда моделирования представляет из себя два программных блока.

Первый блок создает сущности транзакций относительно пользовательских настроек, описанных в разделе «Параметризация сети». На момент появления сущности транзакции известен ее владелец или отправитель и виртуальное время появления транзакции в сети.

Второй блок отвечает за формирование сети. В общих чертах процесс формирования выглядит так:

- Созданные сущности транзакций из первого блока сортируются по возрастанию относительно их времени появления в сети.

- Транзакции поочередно отправляются в сеть и формируют направленный ациклический граф, отталкиваясь от своего вредоносного или добросовестного сценария, который формируется исходя из таких параметров, как: наличие вредоносного сценария агента, алгоритма присоединения транзакций в сеть и задержки сети.

В качестве результата среда моделирования предоставляет модель, состоящую из двух элементов. Первый элемент включает в себя схему графа, где вершины являются транзакциями, а ребра – связями между валидаторами и проверяемыми узлами. А второй элемент представляет из себя таблицу значений кумулятивных весов транзакций для дальнейшей аналитики.

Параметризация сети

Базовая параметризация сети [3] включает в себя описание агентов сети таких, как:

- Количество агентов. Оказывает влияние на распределение транзакций между агентами и, соответственно, на вероятность валидации чужих транзакций.

- Временная задержка между агентами. Влияет на область видимости агентов. Чем выше это значение, тем ниже актуальность сети для агента в данный момент времени.

- Частота отправки транзакций между агентами. От этого значения зависит плотность и масштабируемость сети на протяжении всего эксперимента. Также с помощью этого значения можно моделировать общую активность агентов.

Одним из самых важных параметров конфигурации Tangle является алгоритм выбора узлов, которые будет проверять агент, отправивший транзакцию [4].

– МСМС. Выбор проверяемого узла осуществляется за счет построения пути от самой первой доступной транзакции до транзакции без валидаторов. Выбор между узлами на пути осуществляется случайным образом с помощью вероятностной функции: чем больше кумулятивных голосов у узла, тем выше вероятность того, что этот узел будет выбран. Регулировка вероятностной функции осуществляется с помощью параметра α . Он представляет из себя дробное число выше нуля. Чем выше его значение, тем меньше вероятность того, что транзакция с меньшим количеством голосов будет выбрана.

– Выбор проверяемого узла осуществляется случайным образом среди узлов без валидаторов из области видимости агента.

Также можно настроить количество и метод выбора блужданий: те, которые пришли раньше других или наоборот.

Вредоносные сценарии

На данный момент можно выделить следующие вредоносные сценарии над алгоритмом консенсуса Tangle [5]:

- Атака паразитирующей цепью.
- Разделяющая атака.
- Атака большим весом.

Важно отметить, что атака паразитирующей цепью является самой эффективной в неподготовленных сетях.

Эксперименты

Проведение экспериментов можно разделить на два этапа: построение моделей Tangle с помощью параметризации сети и валидация конфигурации сети на вредоносных сценариях.

В рамках данной работы были проведены следующие эксперименты под влиянием вредоносных сценариев:

- Поведение агентов с задержкой сети.
- Сравнение алгоритмов выбора неподтвержденных узлов.
- Сравнение вероятностных функций блужданий.

Однако самым тонким моментом является подбор вероятностной функции и подбор параметра α , так как во многом от него зависит вероятность успеха атаки паразитирующей цепью или вероятность того, что «ленивые» узлы будут включены в сеть.

Как уже известно: чем меньше α , тем более распределенная сеть получится, то есть тем меньше будет отличаться разница вероятностей между узлом с большим и меньшим количеством кумулятивного веса. Вследствие чего, паразитирующей цепи будет проще внедриться в структуру сети. Так, например, на начальном этапе развития сети при рекомендованном значении $\alpha=0.01$, паразитирующая цепь запросто встраивается в общую сеть и набирает больше кол-во валидаторов, чем оригинальная транзакция, как представлено на рисунке 1. При этом, если

провести тот же эксперимент с более высокими значениями α , то можно получить адекватный и ожидаемый результат, когда паразитирующая цепь будет полностью отклонена сетью, что представлено на рисунке 2.

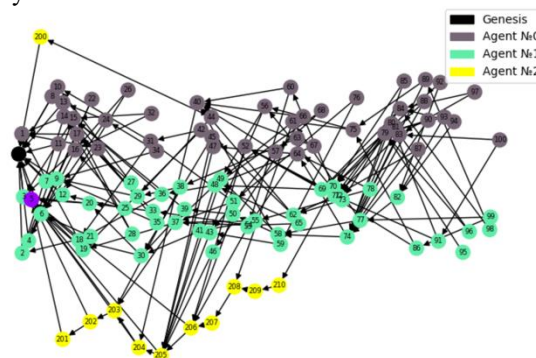


Рис. 1. Успешная атака двойной траты при вредоносном сценарии паразитирующей цепи и низком $\alpha=0.01$

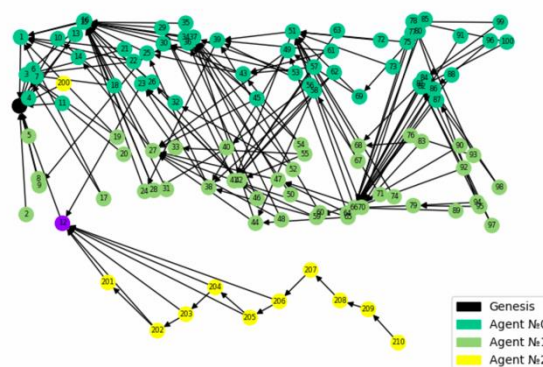


Рис. 2. Неудачная атака двойной траты при вредоносном сценарии паразитирующей цепи и высоком $\alpha=0.5$

Но при этом видно, что от такой политики страдают некоторые добросовестные транзакции, остающиеся без валидаторов.

Проведя некоторое количество опытов, можно прийти к выводу, что лучшим потенциальным решением было бы регулировать данный параметр динамически относительно средней нагрузки на сеть за последнее время. Таким образом α не будет экстремально низким, а высокие значения будут компенсироваться высокой активностью агентов.

Так или иначе, параметр α и вероятностная функция – не единственная возможность борьбы с атакой паразитирующей цепью. Интересным решением является запуск большего количества блужданий и выбор таких, которые проходят наиболее длинные пути. Таким образом, при условии, что вредоносным агентам принадлежит меньшее количество вычислительной мощности сети, чем всем остальным, провести успешную атаку паразитирующей цепью будет очень сложно. С результатом можно ознакомиться на рисунке 3.

Важно понимать, что каждое блуждание подразумевает под собой выполнение вычислений, что замедлит среднюю обработку транзакций в сети в

несколько раз. Если только не заменить точку старта блужданий от самой первой транзакции, до какой-либо временной границы, что также можно регулировать внутри данной среды моделирования.

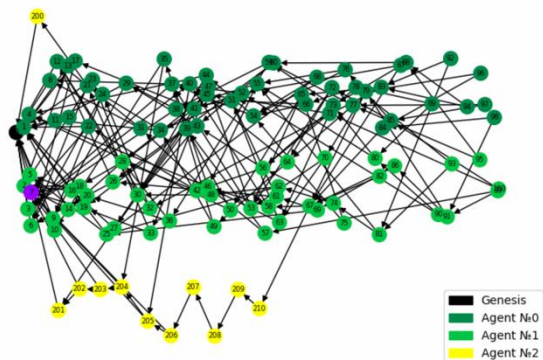


Рис. 3. Неудачная атака двойной траты при вредоносном сценарии паразитирующей цепи, низком $\alpha=0.01$ и группе блужданий

Заключение

Как уже говорилось выше, потенциал среды моделирования позволяет проводить большее количество разных экспериментов, что дает возможность построения всевозможных моделей систем распределенного реестра на основе направленного ациклического графа.

Результатом каждой модели является таблица с соотношением номера транзакции и ее веса, на основе которой можно провести подробный анализ и оценить модель.

Исходя из описанного эксперимента, можно сделать вывод, что поведение Tangle сильно зависит от нагрузки общей сети, что в стандартной конфигурации может подойти не для каждого решения.

В перспективе на основе данной среды можно оценивать возможность применения алгоритмов шифрования [6] методом сравнения их устойчивости к атакам относительно общих задержек.

Список литературы

1. Статья новостного ресурса coincapmarket // Режим доступа: <https://coinmarketcap.com/currencies/iota/> (дата обращения: 09.02.2023).
2. Serguei Popov The Tangle // Режим доступа: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUaU6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf (дата обращения: 09.02.2023).
Научная электронная библиотека ARXIV.
3. Zander M. Simulation of DAG based distributed ledger // Режим доступа: <https://eprint.iacr.org/2018/1062.pdf> (дата обращения: 09.02.2023).
Научная электронная библиотека EPRINT.
4. Bartosz Kusmierz, William Sanders, Andreas Penzkofer, Angelo Caposelle and Alon Gal Properties of the Tangle for uniform random and random walk tip selection // Режим доступа: <https://arxiv.org/pdf/2001.07734.pdf> (дата обращения: 09.02.2023).

09.02.2023). Научная электронная библиотека ARXIV.

5. Ikram Ullah, Gerard de Roode, Nirvana Meratnia and Paul Havinga Threat Modeling—How to Visualize Attacks on IOTA // Режим доступа: <https://www.mdpi.com/1424-8220/21/5/1834/pdf> (дата обращения: 09.02.2023). Научная электронная библиотека MDPI.

6. Lightweight cryptography // Режим доступа: <https://csrc.nist.gov/projects/lightweight-cryptography> (дата обращения: 09.02.2023). Национальный институт стандартов и технологий NIST.

РАЗРАБОТКА СПОСОБА СОЗДАНИЯ 3Д ФОНА ДЛЯ СТУДИИ С ХРОМАКЕЕМ

Заякина А.М.¹, Гронин Д.С.²

Национальный исследовательский университет
«Высшая школа экономики»,

¹департамент компьютерной инженерии,

²департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматриваются способы интеграции интерактивной трехмерной графики в RTSP поток для использования на хромакее в условиях студийной трансляции в реальном времени. Данная разработка проходит в рамках создания приложения-плагина для OBS.

Введение

В сфере видео технологий за последние годы сильно возросла актуальность прямых трансляций с использованием технологии замещения заднего фона. Однако иногда формат трансляции подразумевает динамичный кадр, что вызывает ряд сложностей с наложением заднего фона. Одним из классических решений для записи видео является использование зеленого фона, называемого хромакеем [1]. Поэтому перед выходом продукта съемки в эфир на постмонтаже накладывается графический фон. Во время многокамерной съемки при выборе фона необходимо учитывать перспективу, крупность объектов в кадре. Без учета данных параметров будет видна искусственность картинки, что полностью обесценивает использование хромакея. Использование технологий машинного зрения позволяет автоматизировать работу оператора. Для исключения постмонтажа и возможности проведения трансляции в реальном времени необходимо учесть изменение заднего фона. Это необходимо для того, чтобы при плавном перемещении крупного кадра, в центре которого находится человек, фон и прочие виртуальные модели оставались статичными относительно реальных стационарных объектов. Чтобы упростить подготовку фонов, нужно сделать 3D сцену по размерам студии и при помощи виртуальных камер получать изображения фона в соответствии с ракурсом и крупностью плана

на камерах. В полной версии нужно считывать данные с реальных камер и генерировать для них фоновую картинку. В минимальной версии настройка виртуальных камер производится вручную, картинки из виртуальных потоков подгружаются в OBS, программу для проведения онлайн трансляций, вручную.

Существующие способы

На данный момент классическим подходом к съемкам с хромакеем является обработка после завершения съемки. Для этого используются различные программы монтажа. Однако эти программы не предназначены для работы с видео по ходу трансляции, в режиме online. Работа в них возможно только после завершения записи, что исключает использование этого решения во время трансляции.

Самой популярной программой для стриминга на данный момент является OBS Studio, это свободная программа с открытым исходным кодом для записи видео и потокового вещания. Ввиду ее популярности, существует обширный выбор плагинов и надстроек, значительно расширяющих функционал OBS. Однако, общий спектр открывающихся возможностей не может в полной мере удовлетворить изначальным требованиям. Это связано с тем, что даже при использовании виртуальных камер, невозможно создать динамичную конфигурацию настроек и программных модулей. Рассматриваемая программа работает, опираясь исключительно на сцены и кадры, соответственно при статичной сцене, можно добавить трехмерный фон на хромакей и прочие объекты дополненной реальности. В случае, если ведется трансляция с использованием нескольких предварительных настроек положений кадра, они могут быть настроены таким образом, что при смене кадра будет создаваться эффект смены угла зрения, но это достаточно тонкая и сложная ручная настройка, которая все ещё не позволяет двигать камеру в эфире, например в случае, когда человек находится вне предварительно заданного кадра, что заставляет режиссера резко давать общий план и возвращать исходную настройку положения кадра в течении выступления спикера.

Стоит отметить, что также существуют более серьезные плагины, в основе которых лежат технологии искусственного интеллекта и машинного обучения, опуская то, что в большинстве случаев эти программы являются платными кроссплатформенными нейросетями, они нацелены на видоизменение или замещение части видео и аудио потока путем искусственной генерации или наложения слоев. Однако нейросети почти не работают с фоном. Они либо оставляют его без изменений, либо делают монохромным или размытым, создавая эффект расфокусировки, что также не подходит для описанной задачи. Тем не менее, эти функции могут стать весьма полезным дополнением к дальнейшему развитию предложенного ниже программного метода в качестве инструментов обработки видеопотока.

Преимущества и недостатки существующих способов

Обзор наглядно показывает отсутствие в свободном доступе простого и надежного решения поставленной задачи. Любой из существующих способов имеет один или несколько недостатков из следующего перечня: нет потокового воспроизведения видеоданных в реальном времени, нестабильность, плохой композитинг, то есть совмещение двух и более слоев, когда в качестве маски используется цвет или яркость цвета, исходная мультизадачность, приводящая к плохой адаптации для реализации поставленной задачи, отсутствие автономности и дороговизна.

Предлагаемая технология

Для разработки оптимального метода решения поставленной задачи, был взят во внимание следующий ряд обязательных критериев:

1. Метод должен являться концептуально новым решением с точки зрения подхода к студийным съемкам;

2. Разработанная технология должна открывать обширный спектр возможностей, для дальнейших разработок на базе предложенного метода, а также позволить значительно сократить затрачиваемые человеческие и временные ресурсы при выполнении производственных задач студии;

3. Итоговое приложение должно быть автономным и экономически независимым, а именно, в работе оно не должно опираться на ресурсы и вычислительные мощности сторонних web- или платных сервисов.

За основу была взята игровая среда разработки Unity 3D, которая является современным кроссплатформенным движком для создания игр и приложений, разработанный Unity Technologies. Выбор основан на обширном функционале базовой версии программы, позволяющем:

1. Воссоздать реальную студию в виртуальном трехмерном пространстве

2. Координировать виртуальные и реальные камеры, взаимодействовать с объектами (AR & VR)

3. Делать интеграции с Unity MARS, а именно создавать поверх среды AR Foundation окружение, разработанное в среде Unity MARS, что позволит реализовать AR-плагин, способный добавить функционал интерактивного взаимодействия реальных и виртуальных физических тел внутри одного Real-Time Streaming Protocol (RTSP) потока [2].

Для автономности программы оптимальным решением было отказаться от сращивания двух и более потоков, лучшим инструментом оказались простые C# интеграторы на подобии сопрограмм. Они позволяют обрабатывать RTSP поток в режиме реального времени внутри разрабатываемого приложения.

Создание проектного решения для студии с нуля будет выглядеть следующим образом:

1. Создание виртуальной экспозиции объектов в рабочем пространстве Unity 3D (рис.1)

2. Настройка виртуальных камер и источников света внутри AR-пространства

3. Создание слоев и текстур трехмерных объектов
4. Перезапись текстурных объектов путем замещения существующих текстур стандартной библиотеки Texture 2D на загруженные заранее изображения в формате JPEG или PNG
5. Выполнение внутреннего наложения подвижных фонов из папки StreamingAssets
6. Воспроизведение наложенных фонов через функцию PlayMovie
7. Вызываем и встраиваем воспроизводимый фон функцией Application.dataPath + “/StreamingAssets”[3]
8. Инициализация узлов ввода XR
- 8.1. Описание всех требуемых узлов отслеживания, для поставленной задачи - узел отслеживания Hardware Tracker
9. Настройка устройства отслеживания сцены путем добавления входного устройства “Camera”
10. Интеграция потока RTSP
 - 10.1. воспроизведение потока RTSP внутри сцены
 - 10.1.1. Декодирование программой ffmpeg
 - 10.1.2. Передача потокового мультимедиа при помощи библиотек Live555
 11. Захват видео потока при помощи кадрового буфера
 - 11.1. Добавление объекта FrameBuffer
 - 11.2. Добавление размеров кадра в конструктор Capturing
 - 11.3. Импорт параметров во FrameBuffer
 - 11.4. Реализация метода OnPreRender - вызывается перед началом рендеринга сцены
 - 11.5. Переключение на FrameBuffer - весь рендеринг производится на текстуру, закрепленную за FrameBuffer
 - 11.6. Реализация метода OnPostRender - вызывается после окончания рендеринга кадра
 - 11.7. Запуск ожидания получения готового кадра
 - 11.8. Отключение FrameBuffer
 - 11.9. Копирование текстуры на экран средствами MediaPack, реализованном на основе метода EndCaptureFrame (экспорт кадров может быть реализован иначе, все зависит от задач студии)
 - 11.10. Для контроля скорости передачи добавляем классический счетчик FPS и временной синхронизатор (скорость воспроизведения кадра, а соответственно рендеринга и воспроизведения следующего должна не превышать 40-50 ms)

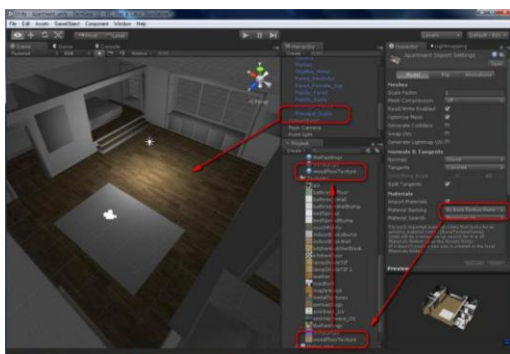


Рис. 1. Пример воссозданной студии

Таким образом, получается приложение, адаптированное под конкретную конфигурацию студии с возможностью редактирования, которое позволяет обрабатывать RTSP поток в реальном времени и покадрово его возвращать после рендеринга. Оно имеет большой потенциал для расширения возможностей ввиду простой программной архитектуры и доступности интеграции нейросетей в процесс обработки потока. Также нейросетевую обработку можно добавлять как плагин, работающий с RTSP потоком до его интеграции в Unity 3D. Дополнительно стоит отметить, что сервисы Unity 3D позволяют добавить интеграцию Unity Mars для реализации более сложных графических эффектов, управляемых практически любыми привязками. А именно наложение слоев в реальном времени внутри потока можно назначать на любые физические и виртуальные объекты. Все это в перспективе позволяет проводить съемки и трансляции любой сложности на соответствующем оборудовании и реализовывать сверхсложные графические эффекты.

Заключение

В ходе работы был проведен анализ существующих способов реализации наложения фона для динамичного кадра, а также предложен программный метод решения поставленной задачи, отличительными чертами которого являются концептуальная новизна, полнота функционала и базовая коммерческая независимость. Делая обобщающий вывод из обзорной части и предложенного метода, очевидно, что в нише бесплатных решений предлагаемый вариант реализации приложения выигрывает по всем трем параметрам, указанным в критериях “Предлагаемой технологии”. Это означает, что метод имеет право на практическое применение и может стать достойной современной заменой, с существенным потенциалом развития, существующих утилит.

Список литературы

1. Нечаева Л.В., Быков Э.Ю., Павлова К.С. Хромакей в помощь преподавателю / Аллея науки. - 2018. - т. 5. - №5. - с.377-381
2. Gwendal S.; Petrangeli S.; Carr N.; Swaminathan V. Streaming a Sequence of Textures for Adaptive 3D Scene Delivery / 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) - 2019.
3. Streaming assets. URL: <https://docs.unity3d.com/Manual/StreamingAssets.html> (Дата обращения: 15.01.2023)

ОБЗОР ПОДХОДОВ К РАЗРАБОТКЕ МОБИЛЬНЫХ ПРИЛОЖЕНИЙ

Кофман М.О., Лютиков А.С.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе приведен сравнительный анализ подходов разработки мобильных приложений: серверной верстки и нативного способа. При использовании серверной верстки (SDUI) информация о расположении объектов интерфейса мобильного приложения поступает с сервера, обрабатывается на клиентской стороне и отображается на экране. При нативном методе разработки элементы интерфейса программируются с применением фреймворков, в которых файлы разметки размещены локально и распространяются вместе с исполняемым файлом. У каждого из методов есть свои особенности, достоинства и недостатки, которые рассмотрены в работе, а также сделан выбор в пользу одной из технологий для разработки мобильного приложения многомодульной системы Odoo [1] в рамках проекта 19111.

Введение

Мобильные приложения играют важную роль в трафике мобильных устройств. Согласно статистике издания eMarketer [2], более 85% времени использования мобильного устройства приходится на мобильные приложения, а не веб-версии сервисов. Согласно статистике Gartner [3], продажи мобильных устройств растут, поэтому разработка мобильных приложений прибыльна. Таким образом, в современной индустрии наличие мобильного приложения является стандартом, а его отсутствие ограничивает потенциальный рост бизнеса [4]. Все больше фирм разрабатывают приложения для своих сервисов для взаимодействия с пользователями. Наоборот, компания, которая проигнорирует появившийся спрос на мобильные приложения, рискует потерять клиентов и, как следствие, долю рынка. Напрашивается вопрос о наиболее эффективных подходах к разработке мобильных приложений.

В исследовании мы охватили основные тренды в мобильной разработке и подробнее изучили наиболее перспективный, на наш взгляд, из них – подход серверной верстки.

Понятие серверной верстки подразумевает создание всей клиентской логики на сервере. Клиент в данном случае лишь рисует компоненты интерфейса в заданном сервером формате, что соответствует философии тонкого клиента. В отличие от нативной разработки, где большое количество логики отображения элементов интерфейса вычисляется на клиентах во время работы с приложением.

Цель работы

Целью данной работы является выявление лучшего варианта для разработки интерфейса между

серверной версткой и нативным методом разработки. Результат исследования будет применен при работе над проектом 19111, в котором разрабатывается мобильный клиент для многомодульной Customer Relationship Management – CRM [5] и Enterprise Resource Planning – ERP [6] системы Odoo. Проект заключается во взаимодействии с сервером Odoo, получением информации, отображении на экране, обработке пользовательских действий и отправке ответа на сервер. Для обеспечения гибкой разработки и масштабируемости, мобильное приложение создается на основе модульной архитектуры.

Задачи работы

Для определения подходящего метода разработки необходимо рассмотреть достоинства и недостатки каждого из способов. Кроме того, требуется выявить ограничения и сценарии, в которых возможно добиться значительной экономии ресурсов и необходимых этапов для реализации каждого из способов и сравнить с имеющимися условиями. По результатам сравнения можно сделать вывод о наиболее эффективном подходе для использования в проекте.

Сравнение нативного и кроссплатформенного способа разработки

Нативная разработка – это написание кода под две операционные системы: iOS и Android.

Также данный метод подразумевает использование нативных фреймворков. Например, для iOS это UIKit [7] или декларативный фреймворк SwiftUI [8]. Для Android – декларативный фреймворк Jetpack Compose [9].

Сравнение нативного и кроссплатформенного способов разработки является центральной темой исследования [10]. Авторы статьи среди недостатков нативного подхода выделяют большие временные и трудовые затраты, однако отмечают, что конечный результат имеет лучшую производительность.

Другим выводом исследования является заключение, что выбор метода разработки: нативный или кроссплатформенный, зависит от конкретного случая. Отдать предпочтение кроссплатформенной разработке стоит, если приложение только отображает информацию с сервера, или выпустить мобильное приложение нужно как можно скорее. Однако, если работа мобильного приложения связана с интенсивной обработкой данных или требует доступа к низкоуровневому Application Programming Interface (API) [11], например, Bluetooth, авторы рекомендуют разрабатывать приложение нативными способами.

При нативном подходе необходимо поддерживать код для двух операционных систем параллельно, а также отслеживать согласованность интерфейса, чтобы приложение выглядело одинаково. Увеличивается стоимости и время разработки.

Однако можно использовать основу нативного подхода и для разработки приложений, которые будут работать сразу на нескольких платформах. Такой подход называется кроссплатформенным. Отличие от нативного в том, что приложение получается более универсальным с точки зрения долгосрочной

поддержки. В частности, приходится жертвовать учетом особенностей каждой платформы (нативный способ). Кроссплатформенный вариант сокращает ресурсы для реализации: требуется меньше специалистов, меньше времени и меньше средств для разработки.

Кроме того, достаточно важным является фактор трудовых ресурсов: кроссплатформенные технологии в большинстве своем достаточно новые и экспертизой по ним обладает меньшее количество людей. Этот фактор не учитывается в исследовании [10], однако мы находим его важным при разработке приложений в реальном мире.

Для случаев, где стоит цель создать минимальную работоспособную версию как можно быстрее и при этом имеются соответствующие трудовые ресурсы, следует использовать кроссплатформенную разработку, в противном случае нативная разработка будет более результативна в долгосрочной перспективе.

В проекте 19111 разработчиков кроссплатформенных приложений меньше. Этот недостаток критичный для проекта, поэтому из двух вариантов выбран нативный как более эффективный.

Серверная верстка

Еще один подход к созданию пользовательского интерфейса – серверная верстка, так называемый гибкий подход. Его суть в том, что сервер через API сообщает клиенту, какие элементы и с каким контентом отображать. Ответ от сервера – json формат, с описанием логики отображения элементов на экране. Чтобы внести изменения в интерфейс, достаточно изменить ответ сервера. Таким образом, данный подход делает приложение более реактивным, позволяя вносить изменения в пользовательский интерфейс без выпуска новых версий приложений. Это является плюсом и отличает серверную верстку от других методов разработки интерфейсов. Данный подход возможно реализовать на iOS, Android и в Web, что делает его популярным.

О высоком интересе к технологии свидетельствуют разработки зарубежных и отечественных компаний в сфере серверной верстки. В 2022 году компания Яндекс опубликовала фреймворк DivKit [12] для серверной верстки, который обширно применяется внутри компании.

Сравнение SDUI и нативного подхода

Данная тема была затронута в статье [16]. В качестве преимуществ SDUI отмечена возможность быстрого обновления приложения, не полагаясь на пользователя, возможность использования общей кодовой базы для обеих платформ, большая вариативность доступных для разработки методов за счет создания библиотеки визуальных элементов. В числе недостатков авторы статьи [16] упоминают сформированные пользовательские ожидания от поведения приложения, а также необходимость создания или расширения библиотеки визуальных элементов.

Выводом данной статьи является ряд рекомендаций по использованию SDUI при разработке: эффек-

тивность SDUI падает в случаях, если приложение разрабатывается только для одной платформы. Кроме того, SDUI является не лучшим выбором в случае, если плохо налажена коммуникация между командами разработки смежных компонентов.

У каждого из рассмотренных методов есть свои преимущества и недостатки. Рассмотрим их более подробно.

К плюсам нативного подхода можно отнести, во-первых, более приятный интерфейс: для нативной разработки интерфейса используются гайдлайны, для iOS – Human Interface Guidelines [13], для Android – Material Design [14]. Поэтому дизайн нативного приложения будет более привычен для пользователя. Возвращаясь к сравнению с серверной версткой, существенным преимуществом нативного метода является отсутствие необходимого доступа в Интернет. К тому же нет необходимости на постоянной основе поддерживать сервер, поэтому затраты на инфраструктуру ниже, чем при использовании серверной верстки. И наконец, нативный подход является широко распространенным, поэтому компании проще найти специалистов, работающих в рамках этого подхода, нежели тех, кто сможет использовать серверную верстку. Это также сказывается на общем уровне доступности для фирм.

Из недостатков нативного подхода можно перечислить более низкую скорость загрузки изменений, что происходит из-за локального хранения файлов верстки и необходимости исполнения клиентской логики. Кроме того, по сравнению с серверной версткой, введение новшеств и исправлений ошибок требует обновления приложения.

Серверная верстка в некоторых сценариях дает неоспоримые преимущества по сравнению с нативным подходом.

Во-первых, из-за снижения нагрузки на сторону клиента считается, что подход серверной верстки помогает сократить общее время загрузки изменений. Можно быстрее решать возникающие проблемы и поставлять эти исправления клиенту. Теперь этот процесс не зависит от обновлений приложения, в отличие от нативного подхода. Во-вторых, при серверной верстке мобильные приложения имеют общую кодовую базу. Значит, изменения в этой базе будут распространяться сразу на все подключенные платформы, что также повышает эффективность и быстроту введения новшеств. Это касается, как нового функционала, так и изменения старого.

Однако данный подход также имеет ряд недостатков. Самый явный из них – необходимость доступа в Интернет для любых операций с приложением. Кроме того, при неправильной реализации эффективность и скорость работы приложения может значительно снизиться по сравнению с нативными способами верстки приложения. Помимо этого, использование серверной верстки требует выстраивания нескольких слоев инфраструктуры, которые бывает сложно встроить в существующие решения, особенно если актуальная кодовая база не задокументирована, что случается достаточно часто в коммерческих приложениях. Следовательно, это умень-

шает доступность подхода для малого бизнеса, и с инфраструктурной, и с ценовой точки зрения.

Сравним два подхода, описанные выше, на основе следующих критериев. Рассмотрим поддержку библиотек сообществ, так как для быстрой разработки необходимо наличие развитых библиотек и фреймворков; скорость выпуска обновлений, в проекте 19111 разрабатывается минимально жизнеспособная версия продукта – Minimal Viable Product (MVP), а затем производятся улучшения, которые необходимо доставлять пользователю в короткие сроки; наличие постоянного интернет-соединения, поскольку некоторые данные в мобильном приложении закешированы и часть информации возможно

просматривать оффлайн, то зависимость от подключения к сети является важным критерием для каждого подхода; сложность поддержки существующих решений, так как для поддержания серверной верстки в случае существования уже реализованной серверной части потребуется создать проксирующий сервер, который будет реализовывать клиентскую логику и верстку.

В результате нашего обзора мы составили сравнительную таблицу подходов и критериев для оценки целесообразности использования серверной и нативной верстки.

Таблица 1. Подходы и критерии

Критерий/ Подход	Поддержка библиотек	Скорость выпуска обновлений	Интернет соединение	Сложность поддержки существующих решений
Серверная верстка	меньше библиотек, меньше поддержка сообщества	быстрее, т.к. интерфейс возможно поменять на сервере	постоянное наличие для отображения интерфейса	встраивание в существующие решение сложнее из-за необходимости поддерживать старые контракты и преобразовывать их на серверную верстку
Нативная верстка	поддерживается больше библиотек	дольше, т.к. необходим выпуск новых версий в магазины приложений	не нужно	встраивание в существующие решения не превышает затрат на разработку нового решения

Выбор для разработки

Учитывая проведенный сравнительный анализ, для работы более эффективно выбрать серверную верстку, но для раскрытия ее потенциала требуются определенные условия. Среди этих условий можно назвать отсутствие уже существующей кодовой базы и наличие ресурсов на обеспечение сервера, содержащего верстку клиента. Когда эти условия не могут быть выполнены, эффективность использования серверной верстки падает. Поэтому для наилучшего результата проекта выберем нативный подход.

Заключение

Проведен обзор и сравнение серверной верстки и нативного подхода к разработке пользовательского интерфейса. Каждый метод имеет свои преимущества: для серверной верстки характерна гибкость разрабатываемого приложения, отсутствие привязки к выпускам новых версий для изменения интерфейса, но для стабильной работы необходимо наличие постоянного интернет-соединения. Для нативного способа характерен более удобный интерфейс, так как разработка ведется для обеих платформ, время и стоимость ее увеличивается.

Также рассмотрены условия, в которых каждый из подходов является наиболее эффективным. Проведен анализ применимости подходов в реальных условиях работы над проектом. С учетом всех аспек-

тов выбран наиболее подходящий метод – нативная разработка пользовательского интерфейса.

Список литературы

1. Open source ERP and CRM | ODOO // ODOO URL: https://www.odoo.com/ru_RU (дата обращения: 26.01.2023).
2. US Mobile Time Spent 2020 | Insider Intelligence Trends, Forecasts & Statistics // Insider Intelligence URL: <https://www.insiderintelligence.com/content/us-mobile-time-spent-2020> (дата обращения: 26.01.2023).
3. Gartner Says Global Smartphone Sales Grew 6% in 2021 // Gartner URL: <https://www.gartner.com/en/newsroom/press-releases/2022-03-01-4q21-smartphone-market-share> (дата обращения: 26.01.2023).
4. Michel Ehrenhard, Fons Wijnhoven, Tijs van den Broek, Marc Zinck Stagno Unlocking how start-ups create business value with mobile applications: Development of an App-enabled Business Innovation Cycle // Technological Forecasting and Social Change. - 2017. - №115. - С. 26-36.
5. Что такое CRM-система: виды, задачи, сервисы - Определение | SendPulse // SendPulse URL: https://sendpulse.com/ru/support/glossary/crm#content_s_ubj_matter (дата обращения: 26.01.2023).
6. Что такое ERP | Определение планирования ресурсов предприятия | SAP Insights // SAP Insights

URL: <https://www.sap.com/cis/insights/what-is-erp.html> (дата обращения: 26.01.2023).

7. UIKit | Apple Developer Documentation // Apple Developer Documentation URL: <https://developer.apple.com/documentation/uikit> (дата обращения: 26.01.2023).

8. SwiftUI Overview - Xcode - Apple Developer // Apple Developer URL: <https://developer.apple.com/xcode/swiftui/> (дата обращения: 26.01.2023).

9. Jetpack Compose UI App Development Toolkit - Android Developers // Android Developers URL: <https://developer.android.com/jetpack/compose> (дата обращения: 26.01.2023).

10. Marchuk A. Native Vs Cross-Platform Development: Pros & Cons Revealed // URL: <https://www.uptech.team/blog/native-vs-cross-platform-app-development> (дата обращения: 23.04.2023).

11. Что такое API // URL: <https://habr.com/ru/articles/464261/> (дата обращения: 23.04.2023)

12. Компания Яндекс — Главные новости — Яндекс выложил в open-source DivKit — инструмент для ускорения разработки мобильных приложений // Yandex URL: https://yandex.ru/company/press_releases/2022/26-08-2022 (дата обращения: 26.01.2023).

13. Human Interface Guidelines - Human Interface Guidelines - Design - Apple Developer // Apple Developer URL: <https://developer.apple.com/design/human-interface-guidelines/guidelines/overview/> (дата обращения: 26.01.2023).

14. Material Design // Material Design URL: <https://m3.material.io/> (дата обращения: 26.01.2023).

15. Garbade M. Native vs. cross-platform app development: pros and cons // URL: <https://codeburst.io/native-vs-cross-platform-app-development-pros-and-cons-49f397bb38ac> (дата обращения: 23.04.2023)

16. Bright Digital Server-Driven UI with iOS - What You Need to Know // URL: <https://brightdigit.com/articles/server-driven-ui-ios/> (дата обращения: 23.04.2023)

КУРС СЕТЕВЫЕ ВИДЕОТЕХНОЛОГИИ

Соколовская К.Е.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В последнее время всё чаще обсуждается тема нехватки квалифицированных инженерных кадров. Основной проблемой является не столько нехватка знаний, сколько отсутствие опыта их практического применения в работе у выпускников [1]. В данной статье описан подход к построению практической

работы на примере дисциплины «Сетевые видеотехнологии».

Введение

В Учебной лаборатории сетевых видеотехнологий уже несколько лет разрабатывается обновленная версия курса «Сетевые видеотехнологии», которая позволила бы научить студентов критическому мышлению, работе в команде, решению прикладных задач и внедрению готового продукта в среду исполнения. Этот курс должен развить в обучающихся умение работать на результат, а также снизить для них порог вхождения в проекты по тематическим направлениям работы лаборатории. Сама лаборатория выступает в роли ресурсной базы и внутреннего заказчика. Медиацентр [2], который также является частью лаборатории, — это основная точка внедрения проектных работ. В нём студенты могут проводить эксперименты и делать проекты на стыке классических сетевых технологий с теле- и видео- индустриями. Он является ключевой точкой апробации разработок и поэтому определяет подход к построению курса.

Построение работы в курсе

Для успешного завершения дисциплины в рамках практической части курса и выделенных на дисциплину часов самостоятельной работы студенты должны выполнить проект или написать НИР. Для этого студенты разделяются на команды от одного до трёх человек, и выбирают себе задачу из заранее сформированного списка или предлагают свою идею, которая соответствует тематике курса. Качество и эффективность работы команд оценивается на 6-ти контрольных точках, представляющих стадии развития проекта.

Для оценки результатов работы на большинстве этапов было выбрано пиринговое оценивание работ, в котором учитывалось как экспертное мнение, так и оценки студентов курса. По заранее заданным критериям, студенты должны оценить выступления других команд. Исследования показывают, что результатам этой оценки можно доверять, при выполнении описанных ниже условий [3]. Отклонения от итоговой преподавательской оценки либо несущественно, либо отсутствует вовсе.

Для того чтобы исключить завышение/занижение оценок других команд, для каждого студента подсчитывается коэффициент доверия. Индивидуальная оценка, которую студенты получают за этап, состоит из оценки, получаемой за выступление команды на защите этапа (до 8 баллов из 10) и персонального коэффициента доверия (от 0 до 2 баллов из 10).

Количество критериев может варьироваться, но оценка выставляется по трёхбалльной шкале, где 0 — неудовлетворительная для данного этапа работа, 1 — результат ниже ожидаемого, 2 — ожидаемый результат на данном этапе, 3 — результат выше ожидаемого. Каждый балл по каждому критерию описывается в критериях оценки,

там детально указывается, что должно быть достигнуто, чтобы получить этот балл.

Оценкой за выступления конкретной команды является взвешенное среднее из всех оценок. При этом коэффициент доверия экспертов (преподавателя и ассистентов) всегда имеет вес равный единице, а вес оценки других студентов соответствует их коэффициенту доверия.

Коэффициент доверия пересчитывается каждую контрольную точку, чтобы научить студентов оценивать работы на основе объективных фактов, а не личных предпочтений и учиться не только на своих ошибках, но и анализировать ошибки коллег. Формула расчёта коэффициента доверия для каждого студента: если разница по модулю между оценкой эксперта и оценкой студента меньше или равна двум, то доверие к студенту равно 100%, а его вклад в оценку приравнивается к экспертному и равняется единице. Если разница превышает 2 балла, то за каждый следующий балл студент теряет по 20% от своего максимума доверия. Пример расчёта коэффициента доверия приведён в таблице 1.

Таблица 1. Расчёт коэффициента доверия

	Крит. 1	Крит. 2	Крит. 3	Крит. 4
Эксперт	2	3	2	1
Студент	1	1	1	2
Разница	1	2	1	1
Итоговая разница	5			

Студент оценил команду с расхождением с экспертом на 5 баллов. Это на три балла больше чем допустимая ошибка, поэтому его коэффициент доверия составляет 40%: $1-3 \cdot 0,2 = 0,4$.

У студентов нет ограничения на количество команд, которые они обязаны выслушать. Можно оценить, как только одно выступление, так и все сразу. Коэффициент доверия будет считаться как среднее из всех попыток. За отсутствие хотя бы одной оценки студент получает 0.

Динамика коэффициентов доверия

Поскольку коэффициенты доверия пересчитываются для каждого студента, можно выявить тенденции в их изменении. Основная гипотеза заключалась в том, что со временем студенты начнут лучше сопоставлять выступления товарищей с критериями и выставлять более объективные оценки. К концу года планировалось, что активное большинство попадет в группу «высокого доверия»: коэффициент

доверия выше 80%. Изменения в коэффициенте доверия представлены на рисунке 1.

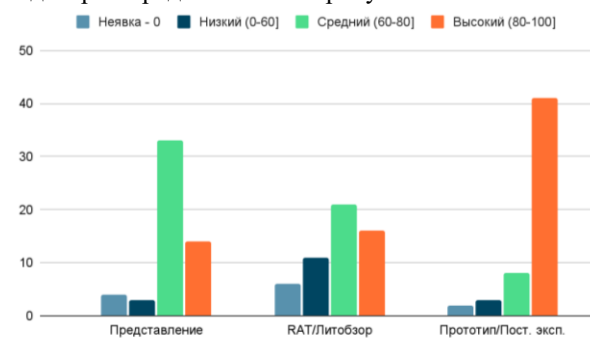


Рис. 1. Количество студентов в каждой из групп «доверия»

Всего в группе 54 студента. Во время представления своих проектов/НИРов больше 60% из них попадали в область среднего коэффициента доверия, а 25% — в область высокого коэффициента. Ко второй контрольной точке (RAT/Литобзор) группа «высокого доверия» почти не изменилась, однако почти треть студентов из области «среднего доверия» перешла в низкую. Это связано с тем, что студентам не было чётко сказано, что коэффициент будут пересчитывать, поэтому они не особо внимательно следили за выступлениями друг друга. К третьей контрольной точке (прототипу/постановке задачи) ситуация нормализовалась, а график принял форму, которую ожидалось увидеть ближе к июню.

На основании имеющихся данных можно предположить, что данная тенденция продолжится. Вероятно, что большинство студентов не ухудшит своё положение, а скорее увеличит коэффициент доверия внутри «высокой» группы. Данный график позволяет заключить, что одна из целей дисциплины, развитие навыков критического мышления и умения оценивать по объективным параметрам, достигается уверенно. Кроме того, в будущем следует заранее сообщить студентам о пересчете коэффициентов доверия, а также акцентировать на этом особое внимание. Это позволит избежать отклонений в росте экспертности.

Анализ успешности проекта

Все критерии, по которым студенты оценивают работы делятся на 4 категории: позиционирование продукта/обоснование исследования, продукт/исследование, документация, фидбек пользователя/рецензия. Такая компоновка необходима, чтобы привести разрозненные критерии оценок к единому виду, по которому можно будет оценить динамику работы команд.

Не все категории начинаются и заканчиваются одновременно. Это связано с тем, что каждая категория описывает конкретные этапы создания работы. График применимости каждой категории по этапам проекта представлен на рисунке 2.

	Представление	RAT/PoC	Прототип	MVP	MUP	Защита
Позиционирование						
Продукт						
Документация						
Фидбек						
	Представление	Литобзор	Постановка экс.	Результат эксп.	Статья	Защита
Обоснование						
Исследование						
Документация						
Рецензия						

Рис. 2. График применимости начала и окончания каждой категории по этапам проекта / исследования

Каждый из типов работ фактически имеет две оценки в конце года: по готовности работы и, по мнению пользователя. Веса этих оценок равны.

Поскольку объёмная работа длиной в семестр может оказаться довольно сложной, необходимо иметь способ выявить проекты, которые близятся к провалу. Для проведения данного анализа была составлена таблица [4] с успешностью категорий для всех проектов и НИР. Рассмотрим самые показательные случаи: успешная работа и неуспешная работа.

Графики могут попадать в разные зоны со следующим значением:

- 1) 2,5-3: соответствует состоянию «отлично», темп работы выше ожидаемого;
- 2) 1,5-2,5: соответствует состоянию «хорошо», работа в нормальном темпе;
- 3) 0,5-1,5: соответствует состоянию «удовлетворительно», проект отстаёт от ожидаемого темпа;
- 4) 0-0,5: соответствует состоянию «неудовлетворительно», проект проваливает какие-то из этапов.

Оттенками синего на графиках обозначены оценки, которые выставляют студенты, оттенками оранжевого — эксперты. При оценке успешности основное внимание уделяется графикам экспертов. Студенческая оценка нужна для анализа отклонения оценки от экспертной.

На рисунке 3 приведён график проекта «SIP intercom». Данный график демонстрирует пример неуспешного проекта. Позиционирование, документация и продукт по мнению экспертов к моменту прототипа сходятся в одной единственной точке — 1. При этом график позиционирования к этому моменту завершён. Столь низкое состояние «позиционирования» на графике означает, что разработчик плохо понимает, как должен работать готовый продукт и для кого он его делает. До защиты этапа 4 проект так и не представил свою разработку заказчику, а также не появился на защите. Поэтому в этапе «MVP» все графики сходятся в 0. Вероятнее всего, к защите этот проект подойдёт в состоянии, которое требовалось бы от MVP, а не от готового продукта.

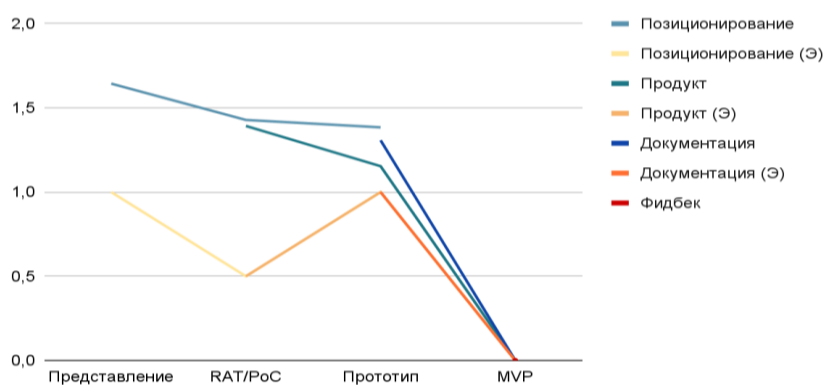


Рис. 3. Неуспешный проект

Пример успешной НИР представлен на рисунке 4. Грамотная работа команды с критериями позволяет ей отчитываться за проделанную работу так, чтобы всегда находиться в отличной зоне. Однако

этот навык будет не так значим, если в результате команда не предоставит пользователю что-нибудь стоящее.

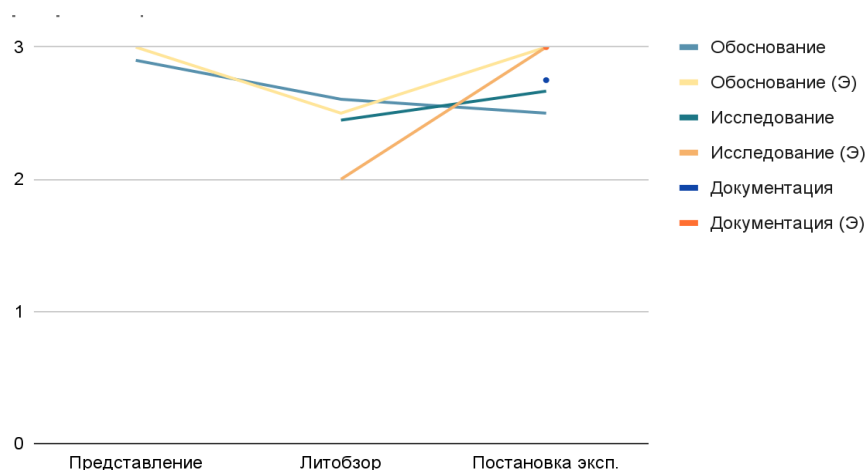


Рис. 4. Успешный проект

В итоге, можно выявить следующие критерии для предсказания успеха проектов/НИР:

Работа будет не успешной, если:

- 1) график «позиционирование»/«обоснование» не превысил и/или планомерно спустился ниже значения 1 к концу;
- 2) все графики находятся ниже значения 1;
- 3) команда пропускает стадию(и) и не пытается успевать за дедлайнами.

Работа будет успешной, если:

- 1) график «позиционирование»/«обоснование» планомерно растёт к значению 3 и/или превышает всё время своего существования превышает 2;
- 2) остальные графики принимают значение большее чем 1,5.

Из полученных данных также следует, что студентам проще оценивать нормальные и успешные проекты. Даже на втором этапе есть места, где значения студентов и экспертов почти сходятся. Вероятно, это связано с тем, что полное соответствие критериям определить проще, чем какие-то недочёты, особенно, если ситуация спорная.

Вывод

В данной статье описан вариант реализации проектной работы на примере курса «Сетевые видеотехнологии». В нём очные занятия не являются пассивным элементом обучения. Данные первой половины курса говорят о том, что система работает достаточно эффективно и позволяет развивать студентам критические и профессиональные навыки.

Список литературы

1. Garousi V. et al. Closing the Gap between Software Engineering Education and Industrial Needs // IEEE Softw. IEEE Computer Society, 2020. Vol. 37, № 2. P. 68–77.
2. Заякина А. и др. Развертывание студенческой телевизионной студии в МИЭМ НИУ ВШЭ // межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов имени Е.В. Арменского. 2021. P. 166–167.
3. Topping K. Peer Assessment Between Students in Colleges and Universities //

<http://dx.doi.org/10.3102/00346543068003249>. Sage PublicationsSage CA: Thousand Oaks, CA, 2016. Vol. 68, № 3. P. 249–276.

4. Сырые данные.xlsx - Google Таблицы [Electronic resource]. URL: <https://docs.google.com/spreadsheets/d/118HIPNdPkjcBbJQTsUa8uSPJcJC0O0Z6/edit#gid=132446127> (accessed: 15.04.2023)

МОДУЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ СЕТЕВЫМИ ВИДЕОКАМЕРАМИ И ПРОГРАММНЫМИ КОМПОНЕНТАМИ ВИДЕОСТУДИЙ С ПОМОЩЬЮ ФИЗИЧЕСКИХ УСТРОЙСТВ ВВОДА

Ефремов А.М., Елисеев А.М.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Данная работа заключается в поиске решения проблемы управления сетевыми видеокameraми и приложениями посредством различных устройств ввода (пультов) в сфере сетевой видеозаписи. В разработку входят задачи по созданию модульной архитектуры сервиса для прослушивания и обработки сигналов с отдельных устройств, использующих различные протоколы передачи данных, а также перенаправления сообщений к конечным сетевым устройствам или программному обеспечению, реализованных согласно шаблону проектирования подписки на события каналов. В случае с управлением сетевыми PTZ видеокameraми, поддерживающими протокол ONVIF, рассматривается проблема временных задержек в управлении их движением, включающая в себя обработку потоков информации на высоконагруженных веб-серверах, которыми являются упомянутые камеры.

Введение

С появлением сетевых видеокамер их использование получило широкое распространение в сфере проведения прямых трансляций, видеоконференций, записи роликов и проведения прочих мероприятий. Мультимедийные пресс-центры стали одним из мест активного внедрения данной технологии.

Такой тип камер характеризуется возможностью удаленного управления по протоколам VISCA [1] и PELCO-D/P [2] посредством цифровых сигналов по интерфейсам RS485 и RS232 [3]. Однако, в работе будет рассмотрен сетевой протокол ONVIF [4], так как он является наиболее применимым в специфике работы сетевых видеостудий. Особенностью данных телестудий является возможность вести распределенную видеосъемку на множестве локаций, подключенных к сети предприятия. Поддержка протокола ONVIF наиболее распространена на рынке сетевых видеокамер и позволяет камерам выступать в качестве веб-сервера, таким образом формируется однородная структура видеопроизводства, в которой управление и получение видеопотоков с камер происходит по одному сетевому интерфейсу на камере. Передача цифровых сигналов управления по протоколам VISCA и PELCO-D/P в данном случае неэффективна, так как менее распространена на рынке сетевых видеокамер, кроме того, требует наличие интерфейсов RS485 или RS232.

Также в рамках работы телестудии возникает потребность управления не только источниками видеопотоков, но и другими устройствами и программными компонентами. К примеру, MIDI-совместимым [5] аудиомикшером X-AIR X18 и программным видеомикшером OBS.

Актуальность

На рынке программного обеспечения в сфере мультимедийных технологий нет единой системы управления сетевыми устройствами, не привязанной к конкретным моделям управляющих и управляемых устройств или инфраструктуре их поставщиков. При записи эфиров и съемках необходим способ объединить устройства управления съемочным процессом. Основное преимущество разрабатываемой системы – в модульности. Данное свойство позволяет дополнять ее поддержкой разных устройств ввода и управляемых устройств или приложений, что значительно сказывается на возможностях масштабирования системы. Таким образом, она будет обеспечивать прослушивание и обработку сигналов с отдельных устройств, использующих различные протоколы передачи данных (MIDI, HID [6], PELCO-D, VISCA, DMX [7]), а также устройствами со своими, нестандартизированными протоколами управления через программные или аппаратные (RS232 или RS485) интерфейсы (матрицы, переключатели входов).

Большинство поставщиков в своих устройствах привязывают компоненты к строго определенным действиям, что не всегда соответствует требованиям отдельной съемки.

Таким образом, потребители вынуждены приобретать дополнительные устройства под новые цели или искать другие способы создания необходимых съемочных условий.

Разрабатываемая система позволяет пользователям самим определить действия для доступных компонентов пультов путем создания конфигурационных файлов, что позволяет увеличить уровень гибкости и упростить работу съемочной группы.

Кроме этого, важным вопросом остается точность управления устройствами, в частности видеокамер. В представленных на рынке программных и аппаратных аналогах остается нерешенным ряд проблем, связанных с работой протокола ONVIF. В протоколе ONVIF используется протокол SOAP для формирования сообщений (запросов), а для транспортировки этих сообщений используется HTTP. Данная особенность протокола позволяет рассматривать сетевые видеокамеры как веб-серверы. Однако, не каждая камера достаточно производительна, чтобы справляться с большими потоками входящих на нее запросов, из-за чего могут возникать негативные эффекты по типу долгих ответов на запрошенные команды или формирования очереди из команд, каждая из которых тормозит выполнение предыдущих. Разрабатываемая система решает данный тип проблем путем применения двух шаблонов проектирования: подписки на события каналов [8] и асинхронной работы HTTP-запросов [9].

Существующие аналоги

Среди аналогов для обеспечения взаимодействия с сетевыми камерами стоит выделить ONVIF Device Manager (ODM). Данный инструмент позволяет сканировать сеть и искать доступные ONVIF-устройства, которыми можно управлять в трех измерениях PTZ (Pan-Tilt-Zoom). Обычно он используется для настройки параметров видеокамер и позволяет управлять поворотными сетевыми камерами с помощью мыши, что неудобно для операторской работы. Также данный продукт серьезно нагружает ресурсы системы и не имеет интеграционного потенциала с другими устройствами и программными компонентами видеостудии.

Альтернативным решением можно считать специализированные аппаратные контроллеры для ONVIF камер, например Anxinshi ASKB320. Они позволяют комфортно вести операторскую работу, но их функционал ограничен работой с видеокамерами по протоколам ONVIF, PELCO-D/P и VISCA, что также сильно ограничивает интеграционный потенциал. Соответственно, в случае, когда появятся новые управляемые устройства или ПО придется приобретать отдельное специализированное устройство управления.

Цели и задачи

Целью разработки является создание единой модульной системы для управления сетевыми устройствами видеозаписи и программными компонентами применяемых в телестудиях, а также исследование зависимости временных задержек при взаимодей-

ствии с сетевыми устройствами от программно-аппаратной реализации систем управления. В том числе, цель текущей работы заключается в изучении способов сокращения задержек при помощи методов асинхронного программирования.

Сетевые видеокамеры, работающие по протоколу ONVIF, перегружаются сетевыми запросами из-за недостатков синхронных программных реализаций протокола. Разрабатываемый сервис решает эту проблему путем ввода модульной системы, созданной в соответствии с шаблоном проектирования подписки на каналы передачи данных. Таким образом создается возможность перевести синхронное управление с физических устройств на асинхронный программный модуль, используя асинхронный сервис передачи данных в каналах, в данном случае СУБД Redis [10].

Кроме того, реализуемая модульность предполагает возможность совмещения функций управления произвольным количеством разнопрофильных устройств и программным обеспечением в рамках одного пульта или набора устройств ввода. Поддерживаемый функционал в данном случае определяется разработчиками и доступен для назначения по каждому компоненту отдельного пульта в редакторе конфигураций. Таким образом, реализация проекта позволит разработать систему, которая даст возможность пользователям, например, управлять сетевыми камерами и событиями в программных компонентах, пользуясь одним пультом. На рис. 1 показана одна из конфигураций режиссерского MIDI пульта, в которой предполагается управление событиями в видеомикшере OBS, а также управление камерами с помощью джойстика. Другим хорошим примером может послужить унификации протоколов управления видеокameraми, реализуемое с помощью прослушивания сигнальных протоколов (PELCO-D/P или VISCA) на управляющем устройстве с последующей интерпретацией сигналов в сетевые ONVIF-команды.

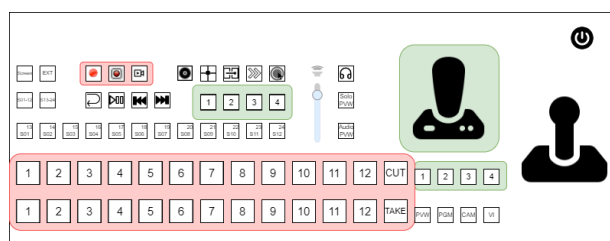


Рис. 1. Расположение органов управления на типовом режиссерском пульте (зеленым – компоненты управления камерами, красным – OBS, без выделения -- другие программы)

Основными задачами в данной работе являются:

1. Изучение способов создания архитектуры программных приложений
2. Создание сервисов для прослушивания сигналов с устройств ввода разных протоколов
3. Создание сервисов для переадресации запросов к сетевым устройствам и программным продуктам

4. Создание архитектуры межсервисной коммуникации

5. Изучение способов сокращения задержек при помощи методов асинхронного программирования

6. Исследование зависимости временных задержек при взаимодействии с сетевыми устройствами от программно-аппаратной реализации систем управления

В системе координации работы студийных пультов с IP-камерами и программными компонентами есть три основных составляющих, которые обеспечивают ее работу.

Сервисы прослушивания сигналов с устройств ввода используются для непрерывной обработки сигналов с пультов. Они содержат логику для сопоставления пользовательских нажатий или прикосновений и действий, относящихся к конкретной клавише или другим компонентам устройства. Информация о привязке действий к компонентам устройств хранится в нереляционной базе данных Redis в виде пар «ключ-значение». Конфигурация устройств происходит путем добавления новых записей или изменения существующих несколькими способами. Например, через редактор таблиц Google Sheets или вручную в соответствии с установленным форматом. При запуске сервиса, происходит обращение к хранилищу по названию соответствующего конфигурационного файла для устройства, который является ключом в базе данных. В процессе работы конфигурационные данные могут быть запрошены из хранилища повторно по запросу других сервисов для обновления. После успешной обработки происходит отправка сообщения следующему сервису, ответственному за переадресацию действия на конечные устройства или сервисы. Также сервисы прослушивания выполняют функцию управления подсветкой для устройств с такой возможностью, путем отправки специальных сообщений обратно на устройство. По принципу модульности для каждого устройства создается отдельный сервис по единому шаблону и разворачивается в системе.

Сервисы переадресации сообщений получают сообщения от сервисов прослушивания сигналов посредством прослушивания соответствующих каналов Redis и создают соответствующее для протокола сообщение. В случае с OBS это будет Websocket [11] событие или запрос, стандартизированные плагином obs-websocket, для сетевых IP-камер – ONVIF запрос. Вместе с этим, по завершению определенных событий для OBS или IP-камер данные сервисы могут отправлять собственные сообщения обратно к сервисам прослушивания сигналов для управления подсветкой.

Принцип работы распределенной системы заключается в обмене сообщениями между отдельными сервисами. Благодаря Redis и его системе создания каналов, а также готовому архитектурному шаблону проектирования [8] подписки на события каналов и передачи сообщений Pub/Sub [12], реализуется система межсервисной коммуникации. Таким образом имеется возможность построить распределенную архитектуру системы и объединить в еди-

ную цепь несколько сервисов, участвующих в обработке сигнала с пульта и доставке сообщения до конечного потребителя. Масштабируемость системы [13] достигается за счет создания нового сервиса и внедрения в общую систему путем подписки на общий Redis канал. В таком случае отправители сообщений, именуемые издателями, не будут напрямую привязаны программным кодом отправки сообщений к подписчикам. С точки зрения количества одновременных подключений разработчиками СУБД предусмотрено ограничение в 10000 клиентов в стандартной конфигурации, что обеспечивает необходимый запас для расширения системы. Тем не менее данная величина доступна для переопределения путем изменения директивы «maxclients» внутри конфигурационного файла. Однако, данное число не должно превышать количество доступных файловых дескрипторов в ОС, иначе предельное значение будет установлено Redis автоматически.

Хранилище Redis одновременно может обслуживать до 232 ключей. Причем, практически каждый из доступных типов данных рассчитан на хранение 232 элементов. Следовательно, с большой долей вероятности можно утверждать, что память хранилища Redis ограничена доступным объемом памяти устройства. В плане расширения текущей системы такое ограничение не представляет большой угрозы, поскольку в базе данных хранятся только записи с конфигурационными данными устройств, число которых во много раз меньше максимального, а объем памяти используемых устройств позволяет сохранять требуемое количество информации.

С целью минимизации времени задержки между действиями по управлению камерами или событиями в видеомикшерах в сервисах переадресации сообщений используется асинхронный принцип работы. Особенно сильно это сказывается на управлении камерой с помощью джойстика, где время отклика является основной метрикой. В таблице 1 приведены данные сравнительного анализа разработанной системы с прямыми аналогами. Были произведено 50 замеров времени отклика устройства при выполнении операции движения (ContinuousMove с векторным параметром Velocity равным 1 хотя бы по одной из осей) и при выполнении операции остановки (ContinuousMove с векторным параметром Velocity равным 0 по всем осям) для каждого из управляющих устройств. Эксперимент проводился в изолированной компьютерной сети, все компоненты (компьютер с сетевым адаптером с пропускной способностью 1 Гб/с и ONVIF Device Manager, аппаратная система управления по протоколу ONVIF Anxinshi ASKB320, Raspberry Pi 4 с развернутой разрабатываемой системой, сетевая видекамера CleverMic 4K 4035UHS) были подключены напрямую в маршрутизатор в интерфейсы с пропускной способностью 1 Гб/с.

Таблица 1. Временные задержки взаимодействия с сетевыми камерами.

Система управления	Среднее время отклика операции движения, мс	Среднее время отклика операции остановки, мс
ONVIF Device Manager	121	217
Аппаратная система управления по протоколу ONVIF	97	104
Разрабатываемая система	125	129

В ходе эксперимента было замечено, что на результат выполнения операции остановки с помощью ONVIF Device Manager влияло время, в течение которого выполнялась операция движения. При анализе передаваемых HTTP-пакетов выяснилось, что ODM непрерывно отправляет запросы движения на камеру, из-за чего на ней скапливается очередь из запросов, на которые надо дать ответ. Камера не успевает синхронно обработать каждый запрос, вследствие чего после отправки запроса на остановку видекамера продолжает выполнять предыдущие запросы.

Проанализировав результаты, можно сказать, что разработанная система выигрывает у программного аналога в задержке операции остановки, благодаря реализации асинхронного обмена сообщениями. Выигрыша перед специализированным аппаратным решением добиться не удалось, однако система обходит его в возможностях интеграций с различными конечными устройствами, не только ONVIF-камерами.

Архитектура системы

В соответствии с проведенным обзором предметной области в сфере разработки программного обеспечения был сделан выбор в пользу микросервисной архитектуры. Ее альтернативой принято считать создание монолитных приложений [14]. В некоторых случаях выбор такого подхода может быть оправдан, но он имеет существенный недостаток в вопросе расширения проектов и внедрения дополнительного функционала из-за роста объема зависимостей между отдельными программными компонентами. Поэтому, в контексте разработки модульной системы управления видеоборудованием было отведено предпочтение именно микросервисной архитектуре с коммуникацией, реализованной на основе брокера сообщений [15]. Сделанный выбор обоснован тем, что для поддержки работы с разными видами устройств и программ требуется учитывать используемые ими протоколы передачи данных и свойства конкретных устройств или программ.

В текущей структуре системы (см. рис. 2) учтены основные требования к масштабируемости и модульности.

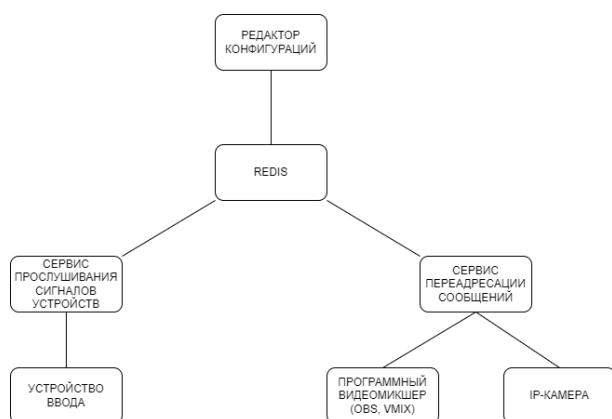


Рис. 2. Упрощенная схема архитектуры системы

Для коммуникации между сервисами используется СУБД Redis, образ которой запущен в виде Docker-контейнера. Она позволяет через специальные библиотеки для языка Python внутри модулей получать доступ к функциям создания информационных каналов коммуникации между сервисами. Одновременно с этим, Redis используется как хранилище конфигурационных файлов в формате JSON, а также сервисной информации. На рис. 3 представлено типовое сообщение, отправляемое сервисом прослушивания сигналов для джойстика. Поле «change» отвечает за выбор типа действия, остальные поля являются параметрами конкретного действия, в данном случае координатами PTZ-движения. Сервис прослушивания сообщений перенаправляет сообщение в требуемый канал в Redis в соответствии с типом действия в поле «change», в данном примере – в канал прослушиваемый сервисом переадресации сообщений для ONVIF-совместимых камер.

```

{
  "change": "ptz",
  "pan": 0,
  "tilt": -0.003905,
  "zoom": 0
}
  
```

Рис. 3. Пример сообщения для управления положением сетевой камеры

Под каждое устройство ввода из комплекта создается собственный Python модуль для обработки сигналов в соответствии с протоколом передачи данных и набором функций. Прослушивание сигналов производится при использовании соответствующих библиотек для языка Python. Каждый модуль запускается в собственном процессе как отдельный сервис.

Для публикации запросов к конечным потребителям в виде ONVIF устройств и OBS разработаны

отдельные Python модули, которые могут обеспечивать передачу информации по протоколам ONVIF и WebSocket соответственно.

Кроме того, предложенная архитектура позволяет провести унификацию способов управления сетевыми устройствами за счет запуска нескольких сервисов прослушивания сигналов с управляющих устройств. Разделение функционала между пультами, работающими по разным протоколам, позволяет создавать конфигурации оборудования под конкретные рабочие сценарии. К примеру, в Медиацентре МИЭМ применяется операторская конфигурация из HID-клавиатуры, MIDI-пульта со световой обратной связью (программируемой подсветкой) и сигнального PELCO-D джойстика. Клавиатура применяется для единичных операций, не требующих визуальной обратной связи (установки камеры в одно из сохраненных положений), MIDI-устройство использует встроенную подсветку для индикации выбора камер и позволяет производить настройку фокуса и экспозиции с помощью энкодеров, джойстик применяется для управления положением камер. Сигналы с каждого из пультов интерпретируются в ONVIF-запросы, посылаемые на камеры. Таким образом в предложенной системе производится унификация работы различных устройств управления, зачастую не предусмотренных для видеопроизводства, реализуется полный потенциал протоколов с большим функционалом. К примеру, функционал сигнального протокола PELCO-D, не предусматривающего настройки экспозиции, дополняется до сетевого протокола ONVIF.

Заключение

Результатом работы является создание единой распределенной модульной системы для управления сетевыми устройствами видеозаписи и программными компонентами, а также её интеграция в инфраструктуру сетевых телестудий.

Список литературы

1. Y. H. Zhao, L. Chai, L. Z. Jin. The Application of PID-DMC Algorithm with Feedforward Compensation in the PTZ Platform. // IOP Conference Series: Materials Science and Engineering. - 2018. - Vol. 428. - No. 1. - p. 012033.
2. O. J. Mokone, et al. Controlling Camera and PDU. - 2015.
3. L. E. Frenzel. RS-485. // Handbook of Serial Communications Interfaces. - 2016.
4. H. W. Di, C. Y. Luo, X. C. Cai. Research and Application of ONVIF Protocol in IP Camera. // Applied Mechanics and Materials. - Trans Tech Publications Ltd. - 2014. - Vol. 568. - p. 1399-1402.
5. H. M. de Oliveira, R. C. de Oliveira. Understanding MIDI: A Painless Tutorial on Midi Format. // arXiv preprint. - 2017.
6. C. Ranta, S. McGowan. Human interface device (hid) profile. // Tilgjengelig på - 2005.
7. O. Luna, D. Torres, R. Americas. DMX512 protocol implementation using MC9S08GT60 8-bit MCU.

// Freescale Semiconductor, Application Note (AN3315). - 2006.

8. R. C. Martin. Design Principles and Design Patterns. // Object Mentor. - 2000. - Vol. 1. - No. 34. - p. 597.

9. D. Namiot, M. Sneps-Sneppé. On micro-services architecture. // International Journal of Open Information Technologies. - 2014. - Vol. 2. - No. 9. - pp. 24-27.

10. M. Paksula. Persisting Objects in Redis Key-Value Database. // University of Helsinki, Department of Computer Science. - 2010. - Vol. 27.

11. I. Fette, A. Melnikov. The WebSocket Protocol. - 2011. - No. rfc6455.

12. Redis Pub/Sub | Redis : [сайт] - 2023. - URL: <https://redis.io/docs/manual/pubsub/> (дата обращения: 15.01.2023). - Текст: электронный.

13. L. G. Williams, C. U. Smith. Web Application Scalability: A Model-Based Approach. // Int. CMG Conference. - 2004. - pp. 215-226.

14. Blinowski G., Ojdowska A., Przybyłek A. Monolithic vs. Microservice Architecture: A Performance and Scalability Evaluation // IEEE Access. 2022. Vol. 10. P. 20357–20374.

15. R. G. Hegde and G. S. Nagaraja. "Low latency message brokers," in Int. Res. J. Eng. Technol., vol. 7, pp. 5, 2020.

СИСТЕМА УДАЛЕННЫХ МЕДИЦИНСКИХ ОСМОТРОВ

*Задорожникова А.А.¹, Зотова А.В.²,
Царегородцева Е.И.²*

*Национальный исследовательский университет
«Высшая школа экономики»,*

¹департамент электронной инженерии,

*²департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе рассматривается проблема, с которой могут сталкиваться водители при прохождении медицинских осмотров (предрейсовых и послерейсовых), а также система, помогающая не только облегчить процедуру, но и значительно сократить время прохождения и решить большинство проблем, связанных с очными медицинскими осмотрами.

Введение

Медицинский осмотр — мероприятие, которое требует вовлечения как минимум одного медицинского работника. Влияние человеческого фактора на результат осмотра (усталость, снижение внимательности, личные проблемы) могут стать причиной губительных последствий, которые могут закончиться летальным исходом. Поэтому проблемы, связанные с проведением медицинских осмотров, например, ограниченная способность медицинских учреждений выполнить все необходимые медицинские

осмотры из-за перегруженности персонала или проблемы с точностью полученных данных в случае недостатка обученного персонала, неправильно выбранных инструментов и/или оборудования, наиболее остро поднимаются в современном мире.

Согласно федеральному закону Российской Федерации «О безопасности дорожного движения», все индивидуальные предприниматели и ООО, связанные с эксплуатацией транспорта, обязаны проводить предрейсовые и послерейсовые медицинские осмотры своих водителей [1]. Это обследование включает в себя следующие процедуры: выявление употребления алкоголя, выявление употребления наркотических и других психоактивных веществ, измерение и оценка артериального давления, измерение и оценка частоты и качественных характеристик пульса, измерение температуры тела, оценка сознания, поведенческой и эмоциональной устойчивости. Приблизительное время этой проверки составляет 15 минут. Это означает, что необходимые медицинские осмотры два раза в день требуют 30 минут рабочего времени медицинского работника и водителя.

Однако все значительно изменилось, когда организации начали использовать системы для удаленных медицинских осмотров [2]. Для оптимизации бизнес-процессов в этой сфере для водителей предлагается решение для проведения данных осмотров удаленно, т.е. водитель самостоятельно проходит осмотр на специальном терминале, на котором установлены все необходимые приборы, а медицинский работник просматривает результаты прохождения и выносит вердикт о его состоянии, используя веб-приложение.

Таким образом, удаленные системы для проведения проверок имеют ряд преимуществ:

1. снижается количество расходов на персонал;
2. практически исчезает человеческий фактор: результаты измерительных приборов автоматически записываются в базу данных, первичный вердикт выносится согласно нормам по каждому показателю, а также за счет встраивания нейронной сети для распознавания лиц, медицинский работник будет уверен, что другой человек не сможет пройти осмотр за водителя;

3. значительно уменьшается время осмотра: теперь длительность процедуры сократилась до 3 минут.

Однако у системы есть некоторые недостатки:

1. Термометрия. Если человек пришел с улицы, термометр может показать температуру ниже, чем она есть на самом деле. Например, 34,5. При этом у медработника не будет возможности измерить показания ртутным термометром на месте.

2. Нужна регулярная проверка медицинских приборов, так как они могут выйти из строя.

3. «Синдром белого халата». Отсутствует возможность выявления терминалом случаев, когда при возникновении у человека волнения из-за прохождения процедуры измерения давления прибор показывает результат выше нормы.

Исследование существующих аналогов

На сегодняшний день на рынке существует несколько аналогичных систем. Например, компании “Медикум” [3], “Tele2med” [4], “MedPoint24” [5] предоставляют возможность прохождения различных дистанционных медицинских осмотров. Среди их клиентов можно встретить “Почта России”, “X5 RETAIL GROUP” и многие другие гиганты своего дела.

Рассматриваемая в данной работе система разрабатывается с набором функций, удовлетворяющих закон Российской Федерации. К отличительным особенностям программного продукта, которые позволяют выделиться на фоне конкурентов, можно отнести разработку необычных решений, таких как авторизация по лицу, система быстрых осмотров, автоматическая генерация статистики и др.

Решение и реализация

На рис. 1 приведена архитектура системы удаленных медицинских осмотров.

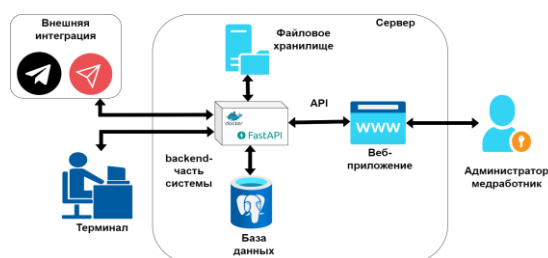


Рис. 1. Архитектура системы удаленных медицинских осмотров

Можно выделить 3 основные части программного обеспечения:

1. веб-приложение для администраторов, диспетчеров и медицинских работников;
2. приложение для терминалов;
3. серверная часть и база данных.

Для разработки клиентской части приложения был выбран Quasar - фреймворк на основе Vue.js [6]. Его преимущество в гибкой библиотеке компонентов, которая значительно ускоряет разработку, и возможности создавать кроссплатформенные приложения.

При создании backend части был выбран FastAPI - фреймворк, написанный на языке программирования Python, помогающий легко и быстро создать API-сервер [7]. Данные хранятся в реляционной базе данных под управлением СУБД PostgreSQL. Решение было принято в пользу этой СУБД, так как она является одной из самых стабильных СУБД и подходит для масштабируемых проектов. Для обеспечения постоянной работы программного комплекса используется средство виртуализации docker. Элементы системы сконфигурированы в docker-compose.yaml, что позволяет надежно и быстро обновлять версии, а также добавлять новые микросервисы и модули.

Приложение для терминала реализовано на базе фреймворка Qt [8] с использованием Python-библиотеки PySide2, для верстки страниц приложения был использован стандартный модуль Qt Quick для QML-приложений, для расширения возможностей приложения и обработки мультимедийного контента был использован модуль Qt Multimedia.

В рамках данной системы водитель проходит медицинский осмотр на терминале. Для этого ему необходимо авторизоваться на нем по табельному номеру. Во время прохождения осмотра ведется видеонаблюдение с фронтальной и боковой камеры, что минимизирует риск подмены человека при прохождении осмотра. Перед осмотром водитель должен внести самостоятельно информацию о жалобах, если они имеются. Далее пользователю необходимо пройти ряд измерений (температура, пульс, давление и алкоголь). Вся информация о пройденном осмотре, включая видеозаписи, отправляется в веб-приложение, где медработник может вынести вердикт о допуске/недопуске водителя. После этого на терминале печатается этикетка, которую водитель вклеивает в свой путевой лист.

При обнаружении алкоголя информация о водителе автоматически передается по SMS в организацию для принятия мер.

Кроме выноса вердиктов у веб-приложения существует еще ряд функций:

1. Возможность просмотра статистических данных по всем организациям и состоянию терминалов;
2. Возможность просмотра статистических данных о проведенных осмотрах;
3. Автоматическая генерация документов (журналов и путевых листов), их просмотр и хранение;
4. Возможность подачи заявок на установку новых терминалов и добавление новых сотрудников;
5. Просмотр настроек терминалов и приборов;
6. Внесение телефонного номера в базу для СМС-информирования.

Для разделения логики взаимодействия с веб-приложением между различными группами пользователей реализована разноуровневая ролевая модель доступа.

Развитие

На рис. 2 можно увидеть, как менялось количество осмотров с марта 2022 года по ноябрь 2022 года.

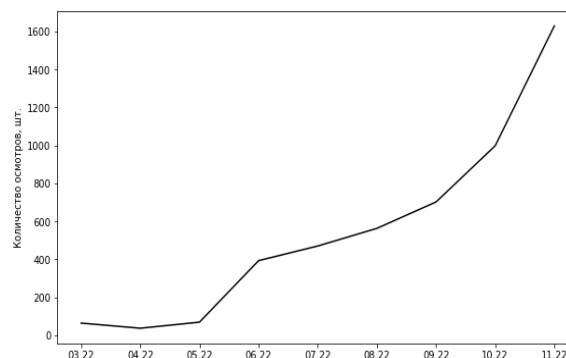


Рис. 2. Рост количества осмотров

На рис. 3 представлен график изменения количества компаний, которые используют в работе данную систему.

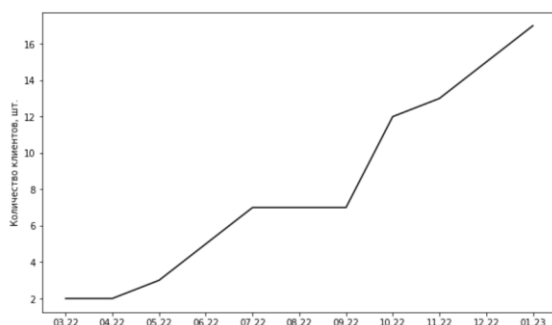


Рис. 3. Рост количества компаний-клиентов

За последние 10 месяцев ежемесячное количество осмотров выросло почти в 26 раз, на текущий момент программно-аппаратный комплекс используется в 17 компаниях для проведения предрейсовых и послерейсовых осмотров. На рисунке видно, что с июля по сентябрь количество компаний оставалось неизменным, однако после произошел резкий скачок. Такому прогрессу способствует непрерывное развитие и доработка новых функций терминала и веб-приложения, таких как: оптимизация основного осмотра на терминале, позволяющая проходить осмотр еще быстрее, разработка новых режимов осмотра, обновление и улучшение интерфейса веб-сервиса для медработников и администраторов и др.

В настоящий момент также ведется разработка программного модуля для биометрической авторизации водителей по лицу, что в будущем почти полностью исключит риски, связанные с подменой пользователя, проходящего осмотр.

Заключение

Таким образом, разработанная система сможет значительно улучшить ситуацию, связанную с прохождением медицинских осмотров для водителей. Данный продукт уже используется во многих компаниях, например, ООО "БлагоСервис" и ГБУ Ритуал. При наличии пожеланий со стороны заказчика ООО "Предрейсовый" или пользователей возможно добавление в систему нового функционала. Благодаря используемому стеку технологий развитие системы не будет вызывать сложностей.

Список литературы

1. Федеральный закон «О безопасности дорожного движения» от 10.12.1995 N 196-ФЗ (последняя редакция) [Электронный ресурс]. URL: http://www.consultant.ru/document/cons_doc_LAW_8585/ (дата обращения: 28.01.2022).
2. Haleem A. и др. Telemedicine for healthcare: Capabilities, features, barriers, and applications // Sensors Int. KeAi Communications Co., 2021. Т. 2.
3. Медуком [Электронный ресурс]. URL: <https://mdcm.ru/> (дата обращения: 28.01.2023).
4. Tele2Med [Электронный ресурс]. URL: <https://tele2med.ru/> (дата обращения: 28.01.2023).
5. MedPoint24 [Электронный ресурс]. URL: <https://medpoint24.ru/> (дата обращения: 28.01.2023).
6. Quasar framework [Электронный ресурс]. URL: <https://quasar.dev/> (дата обращения: 28.01.2023).
7. FastAPI [Электронный ресурс]. URL: <https://fastapi.tiangolo.com/> (дата обращения: 28.01.2023).
8. Qt documentation [Электронный ресурс]. URL: <https://doc.qt.io/> (дата обращения: 28.01.2023).

Секция «Электроника»

ОБЕСПЕЧЕНИЕ УСТОЙЧИВОСТИ К МЕХАНИЧЕСКИМ ВОЗДЕЙСТВИЯМ БОРТОВОГО ЭЛЕКТРОННОГО УСТРОЙСТВА

Пригородов Р.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Работа посвящена проведению исследования вариантов защиты от механических воздействий для сетевого адаптера-зарядного устройства, входящего в состав наземной бортовой аварийно-спасательной радиоэлектронной аппаратуры. Исследование состоит из моделирования воздействий гармонической и случайной вибрации, одиночного и многократного ударов для различных вариантов креплений платы и с использованием виброизоляторов. Расчет коэффициентов нагрузки радиоэлементов позволил сделать вывод об эффективности рассмотренных способов защиты.

Введение

С развитием технического прогресса и усложнением элементной базы устройств, важнейшей задачей при конструировании является обеспечение их устойчивости к механическим воздействиям [1]. Печатные узлы в ходе эксплуатации бортовых устройств испытывают механические нагрузки, и, если они превышают критические значения хотя бы на одном из радиоэлементов, устройства могут выходить из строя или значительно сокращать свой рабочий ресурс [2, 3]. Особенно актуально обеспечение надежности бортового аварийно-спасательного оборудования, один из блоков которого будет рассмотрен в данной работе.

Объектом исследования является сетевой адаптер-зарядное устройство, входящий в состав наземной аварийно-спасательной радиоэлектронной аппаратуры.

Цель работы – повысить устойчивость к механическим воздействиям бортового сетевого адаптера-зарядного устройства путем проведения механического моделирования. В работе будут описаны этапы моделирования и выбор лучшей виброзащиты на основе полученных результатов.

Практическая ценность полученных результатов состоит в том, что благодаря проведенному анализу удалось обеспечить виброустойчивость разрабатываемого бортового устройства.

Моделирование

Для моделирования механических воздействий была выбрана система «АСОНИКА», а именно её подсистемы «АСОНИКА-ТМ» и «АСОНИКА-В».

Подсистема «АСОНИКА-ТМ» позволяет анализировать печатные узлы на влияние тепловых и механических воздействий [4].

Подсистема «АСОНИКА-В» создана для того, чтобы моделировать воздействия на блоки аппаратуры с виброизоляторами и подбирать их, исходя из полученных результатов [5].

Для начала в подсистеме «АСОНИКА-ТМ» была создана модель печатного узла на основе чертежа конструкции. Для каждого радиоэлемента описаны параметры, исходя из их технических условий, созданы точки крепления в количестве 4 штук по краям печатного узла. Уровни механических нагрузок были взяты из ГОСТа 30631-99 «Межгосударственный стандарт. Общие требования к машинам, приборам и другим техническим изделиям в части стойкости к механическому внешнему воздействующим факторам при эксплуатации» [6]. Так как устройство предполагается использовать в наземном транспорте, то нагрузки были выбраны для группы механического исполнения М30, которая определяет воздействия для передвижных безрельсовых комплектных изделий, в том числе для установки на приборных панелях, на которые может передаваться высокочастотная вибрация от двигателя.

При первом варианте расчета было сделано предположение о том, что закрепление печатного узла в блоке жесткое и поэтому механические воздействия на блок будут без изменений передаваться на печатный узел. Моделирование на воздействие гармонической вибрации (диапазон воздействующих частот 200-500 Гц, амплитуда ускорения 5g) показало значительные перегрузки на радиоэлементах печатного узла (рис. 1). На многих элементах значение ускорения превысило 100g, при максимально допустимых 40g.

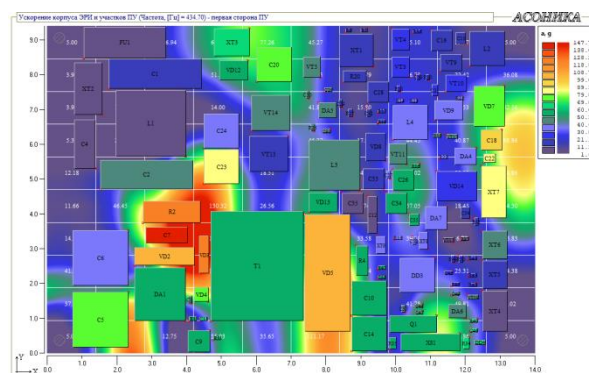


Рис. 1. Ускорение корпусов радиоэлементов и участков ПУ при первоначальном моделировании

При моделировании воздействий случайной вибрации, одиночного удара и многократных ударов, перегрузок не возникло.

Для устранения перегрузок, возникающих при воздействии гармонической вибрации, было принято

решение вначале провести механическое моделирование для различных вариантов точечных креплений печатного узла в блоке (рис. 2), так как при изменении расположения креплений и добавлении их количества возможно достигнуть повышения собственной частоты первого порядка платы примерно в два раза.

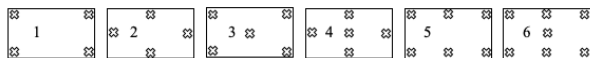


Рис. 2. Варианты размещения креплений на печатной плате

По результатам моделирования для всех способов креплений также были получены перегрузки радиоэлементов при воздействии гармонической вибрации. Некоторые варианты размещения креплений, например - 2, 4, 6, показали перегрузки и при других воздействиях. На основании этого можно сделать вывод о неэффективности обеспечения механической устойчивости таким способом для исследуемого устройства.

Вторым вариантом поиска решения проблемы виброустойчивости стало проведение двухуровневого моделирования, т.е. проведение сначала моделирования блока с добавлением в конструкцию виброизоляторов, а затем полученные в точках крепления печатного узла воздействия задавались при моделировании уже самого печатного узла.

Для моделирования блока была использована подсистема «АСОНИКА-В», в которой была создана 3D-модель корпуса устройства в виде прямоугольного блока (рис. 3).

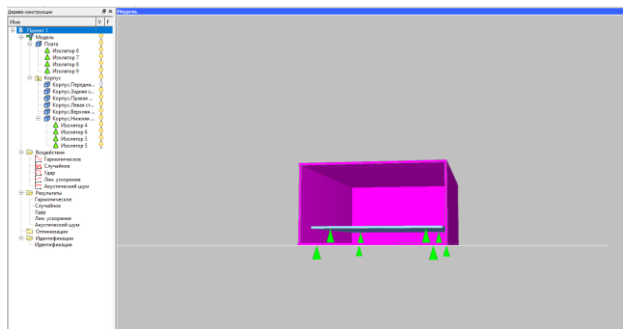


Рис. 3. Модель блока в «АСОНИКА-В»

В основание блока, а также в места креплений на печатном узле были добавлены виброизоляторы. Воздействия на блок были аналогичными, как для первого варианта исследования. Путем подбора разных изоляторов из базы подсистемы «АСОНИКА-В» был выбран виброизолятор АП1-1.8 (рис. 4), который позволил устранить все перегрузки.



Рис. 4. Виброизолятор АП1-1.8

В результате моделирования был получен график зависимости ускорений в точках крепления платы от воздействующей частоты (рис. 5).

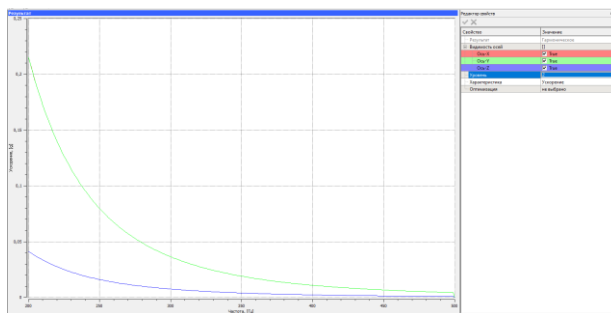


Рис. 5. Зависимость ускорений в точках крепления платы

Полученный график был задан в подсистеме «АСОНИКА-ТМ» в виде графика входных механических воздействий, и было проведено повторное моделирование воздействия гармонической вибрации на печатный узел, исходя из наличия виброизоляторов. Ускорения для гармонической вибрации перестали превышать порог в 10g (рис. 5).

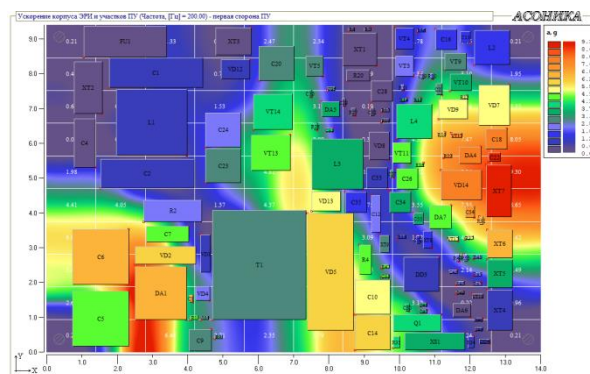


Рис. 6. Ускорение корпусов радиоэлементов и участков ПУ при использовании виброизоляторов

По данным уровней воздействий из ГОСТ 30631-99а было также проведено моделирование и на другие виды нагрузок – случайную вибрацию, одиночный удар и многократный удар для варианта с подобранными виброизоляторами. Результаты показали отсутствие перегрузок.

Анализ результатов

Для анализа полученных данных была использована возможность подсистемы «АСОНИКА-ТМ» выгружать результаты моделирования в виде Excel-таблиц. Для каждого радиоэлемента был рассчитан коэффициент нагрузки. Для всех видов крепления с помощью формул Excel было посчитано количество элементов с коэффициентом механической нагрузки меньше допустимого 0.7, и построены диаграммы, на которых наглядно видно, какой из способов защиты эффективнее (рис. 7). В данном случае – это использование виброизоляторов АП1-1.8, при которых на всех радиоэлементах коэффициент нагрузки не превышает 0.7.

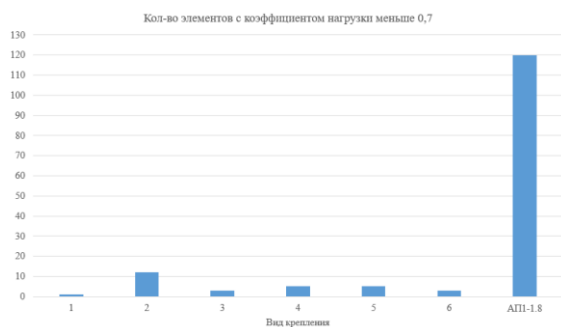


Рис. 7. Диаграмма, показывающая количество элементов с коэффициентом нагрузки меньше 0.7 для разных видов крепления

Развитие

В продолжении исследования будет рассмотрена возможность обеспечения виброустойчивости печатного узла путем добавления ребер жесткости, а также возможность оптимизации параметров виброизоляторов в подсистеме «АСОНИКА-В» для подбора наилучшего решения. Результаты будут сравнены с полученными ранее и сделаны выводы.

Заключение

Приведенное в работе исследование позволяет сделать вывод, что использование виброизоляторов является эффективным методом борьбы с перегрузками на элементах, возникающими из-за механических воздействий. В конкретном разобранным случае они показали себя значительно лучше, чем различные варианты расстановки креплений. Таким образом, можно утверждать, что при использовании виброизоляторов АПН-1.8, сетевой адаптер-зарядное устройство, сможет выдержать оказываемые на него извне механические нагрузки при эксплуатации.

Список литературы

1. Кротова Е.И. Основы конструирования и технологии производства РЭС: учеб. пособие; Ярослав. гос. ун-т им. П. Г. Демидова. – Ярославль: ЯрГУ, 2013. 192 с.
2. Исаев С.А., Столяров О.Г., Торопов В.А. Испытания систем бортового оборудования летательных аппаратов. – Волгоград: ВолГУ, 2017.
3. Kofanov Y.N., Sotnikova S.Y., Subbotin S.A. Method of increasing the reliability of on-board electronic equipment with an analysis of reserves for the electrical, thermal and mechanical loads, in: Proceedings of the 2016 IEEE Conference on Quality Management, Transport and Information Security, Information Technologies (IT&MQ&IS-2016). – St. Petersburg: IEEE, 2016. P. 94-98.
4. АСОНИКА-ТМ – руководство пользователя.
5. АСОНИКА-В – руководство пользователя.
6. ГОСТ 30631-99. Общие требования к машинам, приборам и другим техническим изделиям в части стойкости к механическим внешним воздействующим факторам при эксплуатации: Межгосударственный стандарт: дата введения 28.05.1999 / Межгосударственный Совет по стандартизации, метрологии и сертификации. – Изд. официальное. – Минск: ИПК Издательство стандартов, 1999. 32 с.

РАЗРАБОТКА СТРУКТУРНОЙ СХЕМЫ БЕСКОНТАКТНОГО ИЗМЕРИТЕЛЯ ВИБРАЦИИ

Опекунова А.А., Демидов И.Д.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Целью настоящего исследования является разработка структурной схемы бесконтактного измерителя вибрации, основанного на волоконно-оптическом преобразователе. Для разработки структурной схемы проведено компьютерное моделирование датчика. Адекватность модели подтверждается экспериментом.

Введение

Во время эксплуатации аппараты, используемые в различных отраслях промышленности, могут изнашиваться. Это приводит к развитию дефектов оборудования и возникновению паразитных вибраций и вносит погрешности в процесс производства. Оценить изношенность оборудования и найти причину неполадки можно по параметрам этой дополнительной вибрации. Поэтому возникает потребность в устройствах способных ее детектировать.

Автор статьи [1] приводит основные перспективы развития датчиков вибрации. Например, выполнение измерений в реальном времени, адаптивность к окружающей среде, дистанционное измерение, что обуславливает высокую важность исследований и разработок на эту тему.

Существует два способа измерения параметров вибрации: контактный и бесконтактный. Второй является более удобным. Отмеченным параметрам в большей степени удовлетворяют средства измерения и контроля, основанные на оптическом методе получения измерительной информации.

В [2] проведен обзор основных методов измерения с использованием волоконно-оптических датчиков. Авторами сделан вывод, что, в целом, волоконно-оптические датчики могут обеспечить точность, долговечность и экономичные конфигурации для измерения вибрации. Каждый метод измерения этого типа датчиков имеет достоинства и недостатки и выбор метода зависит от целей применения датчика.

В работе будет рассматриваться волоконно-оптический датчик с внешней модуляцией. Особое внимание будет уделено способу распределения приемных и передающих волокон, а также алгоритму программной обработки результатов. Это позволит с наименьшими затратами добиться высокой точности датчика.

Волоконно-оптические преобразователи

Принцип действия волоконно-оптических преобразователей с внешней модуляцией основан на изменении под действием измеряемой величины интенсивности светового потока при его распространении

в пространстве между передающим и приемным каналами.

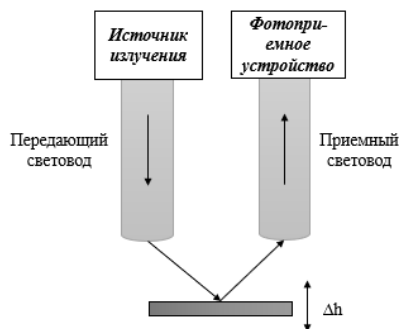


Рис. 1. Структура волоконно-оптического преобразователя с внешней модуляцией

Основными структурными элементами ВОП с внешней модуляцией являются источник излучения, передающий и приемный волоконно-оптические каналы и фотоприемное устройство (рис.1).

Световой поток источника излучения, пройдя по передающему световоду, падает на поверхность объекта измерения. Положение отражающей поверхности определяет значение потока, попавшего в приемный световод, и далее на фотоприемник. Вибрация объекта контроля приводит к появлению переменной составляющей в световом потоке приемника, амплитуда которой пропорциональна виброперемещению поверхности объекта.

Функция преобразования (ФП) ВОП согласно [3] представляет собой сложную зависимость тока фотоприемника от внешних воздействий:

$$I_{\text{фп}} = \sum_{i=0}^N I_i \tau(\varphi)^2 K_{\text{ср}} K_{\text{отр}} K_{\text{шер}} \eta(\lambda) K_{\text{фп}} \quad (1)$$

где $I_{\text{фп}}$ – ток отклика фотоприемника, I_i – ток i -го луча из N вышедших от источника излучения, $\tau(\varphi)$ – коэффициент светопропускания волокна, $K_{\text{ср}}$ – потери при распространении светового потока в среде, $K_{\text{отр}}$ – коэффициент отражения от поверхности, $K_{\text{шер}}$ – коэффициент шероховатости поверхности, $K_{\text{фп}}$ – интегральная чувствительность фотоприемника, $\eta(\lambda)$ – коэффициент спектрального согласования источников и приемников.

Функция преобразования ВОП показана на рис. 2. Она состоит из возрастающего участка от 0 до точки максимума h_{max} , на котором можно выделить участок вблизи точки h_0 , чувствительность датчика на этом участке максимальна. После точки h_{max} функция преобразования начинает убывать. Изменение принимаемого потока практически линейно зависит от изменения расстояния до объекта измерения в h_0 , в то время как вблизи h_{max} практически не зависит, и значит, что рабочая область ВОП — это участок графика функции преобразования от 0 до h_{max} , поскольку характеристика на нем наиболее чувствительна к перемещениям.

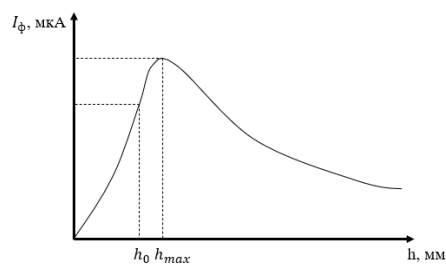


Рис. 2. Функция преобразования ВОП (h_0 – центр участка наибольшей крутизны, h_{max} – точка максимума)

Оценим влияние неинформативных факторов на формировании функции преобразования ВОП.

Наиболее существенными источниками погрешностей в волоконно-оптическом канале являются повреждения волокон: обрыв жилы или трещины.

Большое влияние на отклонение от истинного значения выходного сигнала датчика при бесконтактном измерении имеют внешние неинформативные факторы такие, как свойства отражающей поверхности объекта, внешняя среда, дополнительные колебания руки оператора.

Внешняя среда – источник засветки, так как в разное время суток отличается освещение, также в воздухе могут быть частицы пыли, воды, задымление и т. д.

Предмет исследования может иметь плохие отражающие свойства, например, матовую поверхность или же поверхность с изъянами, трещинами. Уменьшить влияние этих факторов возможно с помощью регулирования яркости источника излучения, например, использование нескольких светодиодов и управление их включением/выключением в зависимости от уровня полученного сигнала на фотоприемном устройстве.

Программная реализации математической модели

С целью проверки правильности полученной математической модели разработана программа, позволяющая построить график ФП от выбранных параметров ВОП: вид светодиода, фотодиода, размеры волокон, характеристики среды и поверхности (рис. 3).

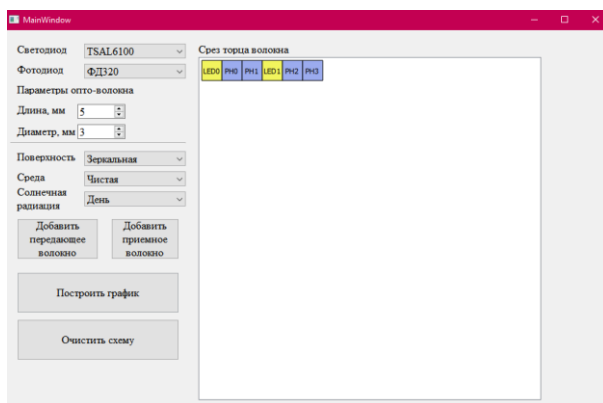


Рис. 3. Главное окно приложения

Расчет происходит в соответствии со следующим алгоритмом. Вначале происходит ввод и выбор параметров конструкции волоконно-оптического канала и данных, характеризующих способ получения измерительной информации. Далее вычисляется диаграмма направленности выбранного светодиода и координаты точки падения луча в соответствии с конструкцией передающего канала и угла выхода этого луча из светодиода. На следующем шаге определяются потери энергии в моменты прохождения светового луча в окружающей среде и его отражения от поверхности измеряемого объекта, проверяются условия попадания элементарного потока на торец приемного световода и возможность его прохождения к фотоприемнику. Затем с учетом потерь в приемном канале рассчитывается вклад прошедшего светового луча в отклик фотодиода. Все перечисленные действия повторяются, пока не будет посчитан вклад N-го луча, вышедшего из светодиода (N - число всех лучей, вышедших из светодиода). В конце расчета результаты выводятся в пользовательский интерфейс.

Разработка структурной схемы прототипа

Характеристики световодов и фотоприемников, как правило, установлены справочными данными, а схемотехническое решение согласующихся устройств и источников питания определяется конкретной метрологической задачей. По этой причине гибкость проектирования ВОП обеспечивается в основном разнообразием конструкторских решений волоконно-оптических каналов. Поэтому исследование влияния конструкции волоконно-оптического канала на функцию преобразования является значимым для разработки ВОП.

Измерение нужно проводить, когда сигнал на фотоприемнике находится в рабочей зоне (рис.2). Но этот участок составляет всего нескольких миллиметров и поддержание необходимого расстояния до объекта измерения достаточно сложно. По этой причине можно использовать несколько фотоприемных устройств.

Расположить их следует так, чтобы точка максимума ФП одного являлась началом зоны чувствительности другого. Таким образом, приближаясь к объекту с некоторого расстояния, проходя точку максимума для одного из фотодиодов, будем оказываться в рабочей зоне датчика. На этапе калибровки необходимо определить на какой дистанции находится эта точка (калибровка представляет собой несколько экспериментов при разных условиях, позволяющих определить необходимые для измерений параметры). В случае если произойдет отклонение (оператор не заметит индикацию о начале измерений и продолжит двигать датчик к объекту), на микроконтроллере переключится канал считывания данных, погрешность измерений не возрастет, так как все еще датчик будет находиться в рабочей зоне.

Влияние внешней засветки, как и влияние отражающих свойств поверхности объекта измерения компенсируются регулировкой яркости светодиода.

Сначала находится точка максимума получаемого сигнала, далее этот уровень сравнивается с тем, что получено ранее экспериментально на этапе калибровки, мощность светодиода увеличивается/уменьшается до того момента, пока измеренное значение не станет равным калибровочному. Только после этого начинается измерение вибрации. При этом никак не повлияет на определение расстояния до объекта и соответственно снова началом измерений, а лишь позволит повысить точность значения виброперемещения.

Основным преимуществом этого алгоритма является его гибкость, то есть в зависимости от текущего положения датчика, оператор может получать сообщения о необходимости придвинуть или отдалить торец световода от поверхности.

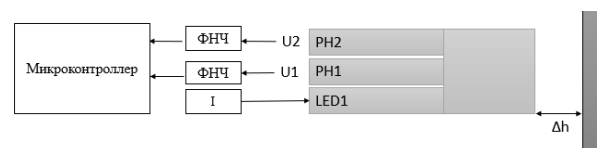


Рис. 4. Структурная схема датчика вибрации

На рис. 4 показана структурная схема измерителя вибрации. Здесь используется трехпучковый световод, обработка данных от которого происходит по описанному выше алгоритму.

Сравнение результатов эксперимента и моделирования

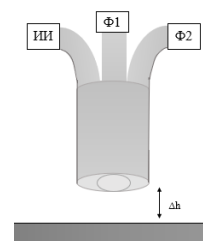


Рис. 5. Структура ВОП

Для подтверждения выдвинутых предположений и проверки адекватности математической модели, проведен эксперимент. Установка для снятия экспериментальной зависимости тока фотоприемника от расстояния до объекта измерения состоит из закрепленной отражающей поверхности, ВОП (3 световода с общим торцом: первый является передающим, остальные – приемные, рис. 5), который закреплен на подвижном элементе, благодаря которому можно проводить юстировку торца относительно пластины, осциллографа и источника тока. Эксперимент проводился при дневном освещении. С учетом этих характеристик также проведено моделирование в нашем приложении. Результаты показали хорошую сходимость в пределах 7 процентов, что подтверждает адекватность математической модели (рис. 6, 7).

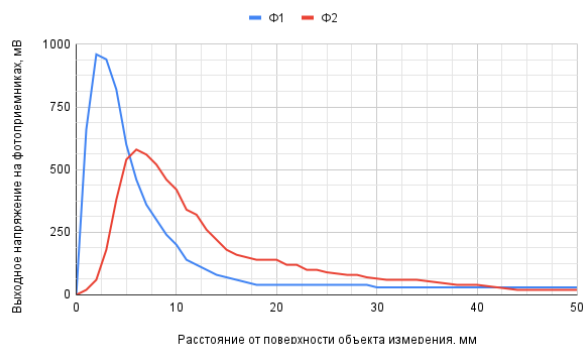


Рис. 6. Результаты эксперимента

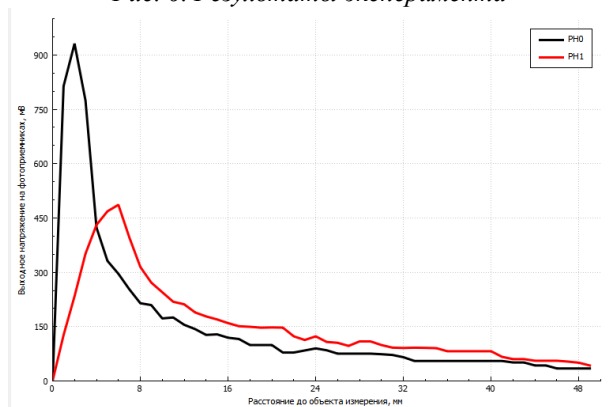


Рис. 7. Результаты моделирования

Моделирование показало, что существует оптимальное соотношение между передающим и приемными световодами, при котором выполняются требования описанного ранее алгоритма.

Заключение

Таким образом, на основе проведенного эксперимента подтверждена адекватность математической модели отклика фотоприемника. С помощью этой программы можно определить конструктивные параметры волоконно-оптического преобразователя, при которых бесконтактный датчик вибрации будет более чувствителен к малым изменениям виброперемещений во время измерений, что определяет ее практическую значимость. Авторами разработан алгоритм и структурная схема, при которых измерения будут проводиться проще и точнее.

Дальнейшие исследования будут направлены на разработку реального прототипа датчика и отладку его работы.

Список литературы

1. Chaurasiya H. Recent Trends of Measurement and Development of Vibration Sensors // International Journal of Computer Science Issues. - 2012. - №9(4). - С. 6.
2. García Y.R., Corres J.M., Goicoechea J. Vibration detection using optical fiber sensors // Journal of Sensors. - 2010. - №2010(4). - С. 12.
3. Зак Е. А. Волоконно-оптические преобразователи с внешней модуляцией. Москва: Энергоатомиздат, 1989. 128 с.

РАЗРАБОТКА МАКРОМОДЕЛЕЙ ОПЕРАЦИОННЫХ УСИЛИТЕЛЕЙ С УЧЕТОМ ВОЗДЕЙСТВИЯ НА НИХ ТЕМПЕРАТУРЫ

Давыдов С.С., Коробейникова Т.И.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В данной статье описана методика построения макромоделей, учитывающих влияние температуры, для операционных усилителей (ОУ) OP191, LM741, 544УД2, К140УД9.

Введение

В настоящее время у компаний, занимающихся разработкой электронных компонентов, есть потребность в получении параметров макромоделей операционных усилителей с учетом температурной зависимости. На данный момент уже созданы макромодели для транзисторов и других электронных компонентов, однако потребность именно в макромоделях операционных усилителей, включающих в себя влияние внешних факторов, осталась до сих пор.

Сейчас разработку такой модели возможно реализовать с помощью программ компьютерного моделирования и компьютерной алгебры.

При тестировании стандартных макромоделей было выявлено, что учет температурного воздействия некорректен по сравнению с документацией к ОУ. Это демонстрируют сравнения графиков коэффициента усиления без обратной связи, полученные в LTspice, с графиками, взятыми из документации к ОУ (рис. 1 – 5).

Целью данной работы является разработка макромоделей ОУ, учитывающих влияние температуры.

Описание предлагаемого решения

Для достижения цели был выполнен ряд задач:

- подбор и анализ литературы по влиянию температуры на параметры и характеристики ОУ;
- подбор и анализ литературы по макромоделям ОУ и по учету в них влияния температуры;
- формирование и верификация SPICE макромодели ОУ по статьям и описаниям ОУ для нормальных условий;
- формирование и верификация SPICE макромодели ОУ по статьям и описаниям ОУ для учета влияния температуры;

Для решения описанных выше задач были проведены тесты в LTspice для макромоделей ОУ OP191 [1], LM741 [2], 544УД2 [3], К140УД9 [4].

Также, для выявления первоначальных условий моделирования и дальнейшего сравнения полученных результатов, была изучена документация по каждому конкретному операционному усилителю. В случае недостатка каких-либо данных в оригинальной документации операционного усилителя была задействована документация существующих анало-

гов (в особенности температурные зависимости коэффициентов усиления ОУ).

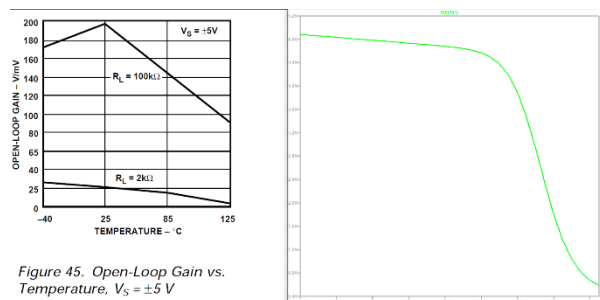


Figure 45. Open-Loop Gain vs. Temperature, $V_S = \pm 5V$

Рис. 1. Сравнение графиков коэффициента усиления без обратной связи из документации для ОУ OP191 и LTspice при синфазном входном напряжении 5 В

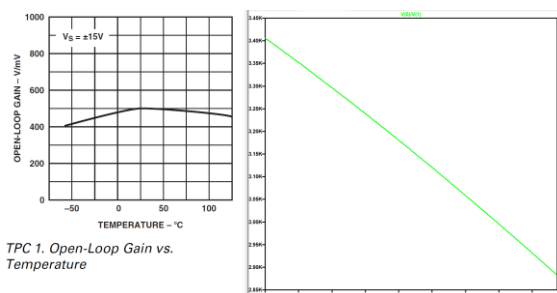


Рис. 2. Сравнение графиков коэффициента усиления без обратной связи из документации для ОУ LM741 и LTspice

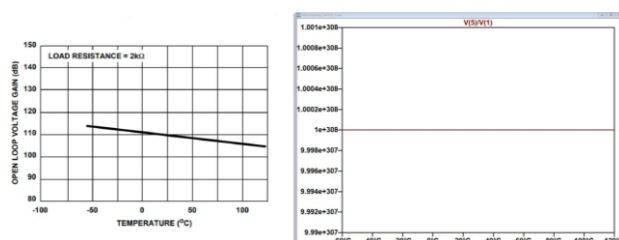


Рис. 3. Сравнение графиков коэффициента усиления без обратной связи из документации для ОУ 544УД и LTspice

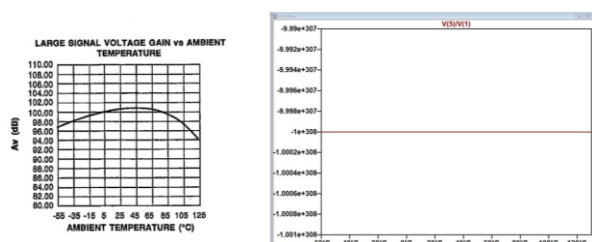


Рис. 4. Сравнение графиков коэффициента усиления без обратной связи из документации для ОУ K140УД9 и LTspice

Макромодели ОУ OP191, LM741, K140УД9 построены на основе статьи [5], так как в ней описывается построение макромодели на основе биполярных транзисторов. Для построения макромодели ОУ 544УД2 была использована статья [6], в которой описан другой пример построения макромодели. В

данных статьях нет учета влияния температуры на характеристики ОУ.

Для упрощения и автоматизации расчетов параметров макромодели ОУ, набор формул из указанных выше статей был реализован в ПО MatLAB (рис. 5).

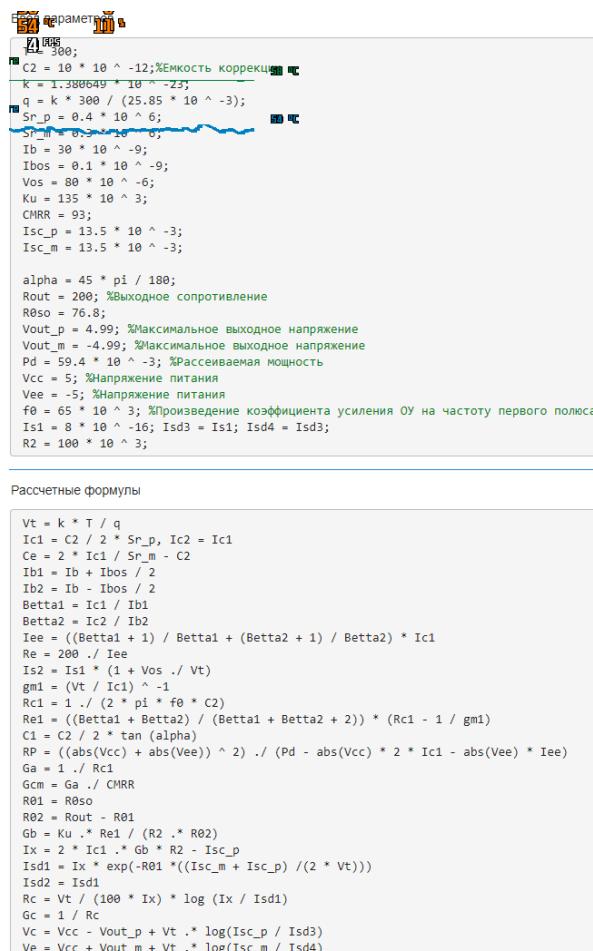


Рис. 5. Реализация набора формул в ПО MatLAB

Расчетные выражение для параметров макромодели ОУ, из которых была составлена математическая модель ОУ в ПО Matlab, были уточнены с учетом влияния температуры. В модели были добавлены зависимости параметров макромодели (рис. 6, 7), таких как G_a , G_{cm} , G_b , G_c от температуры (рис. 8, 9) в соответствии с описаниями исследуемых ОУ OP191, LM741. Код для аппроксимации и построения графиков реализован в ПО MatLAB (рис. 10)

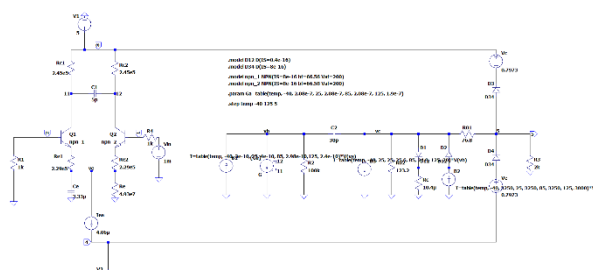


Рис. 6. Электронная схема макромодели ОУ OP191

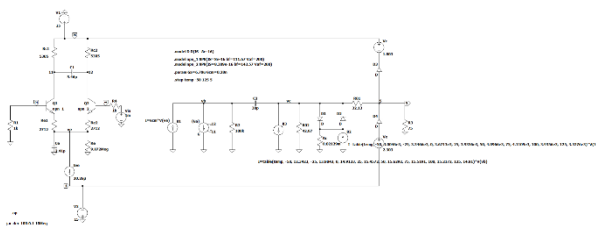


Рис. 7. Электронная схема макромодели ОУ LM741

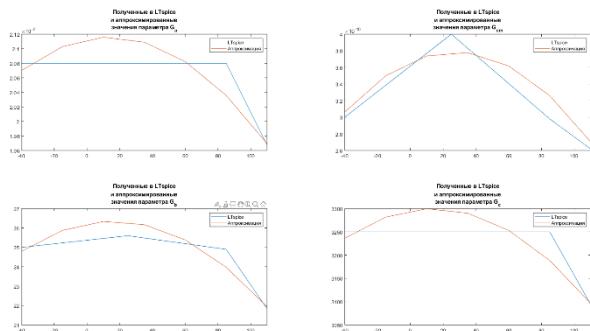


Рис. 8. Графики полученных в LTspice и аппроксимированных значений G_b , G_c , G_a , G_{cm} для ОУ OP191

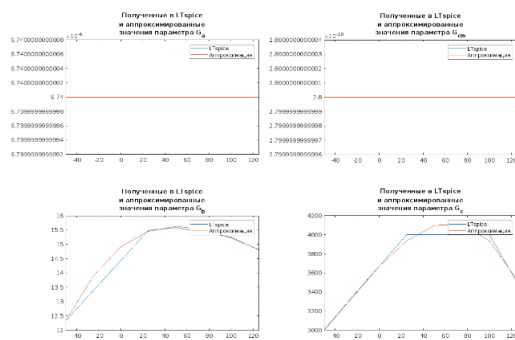


Рис. 9. Графики полученных в LTspice и аппроксимированных значений G_b , G_c , G_a , G_{cm} для ОУ LM74

```
subplot(2, 2, 1)
plot(x, Ga)
hold on
t = -50 : 25 : 125;
[pGa, SGa, MUGa] = polyfit(x, Ga, 3);
Ga1 = polyval(pGa, t, SGa, MUGa);
plot(t, Ga1)
title("Полученные в LTspice")
txt1 = {"и аппроксимированные", "значения параметра G_{a}"};
subtitle(txt1, FontWeight='bold')
legend('LTspice', 'Аппроксимация')
xlim([min(t) max(t)])
hold off
```

Рис. 10. Пример аппроксимации и построения графиков зависимостей параметров макромодели от температуры

Далее в LTspice были промоделированы зависимости коэффициентов усиления от температуры для макромодели ОУ OP191, LM741 (рис. 6, 7). На рисунках 11 и 12 представления сравнения полученных зависимостей, с зависимостями из документации к аналогичным ОУ.

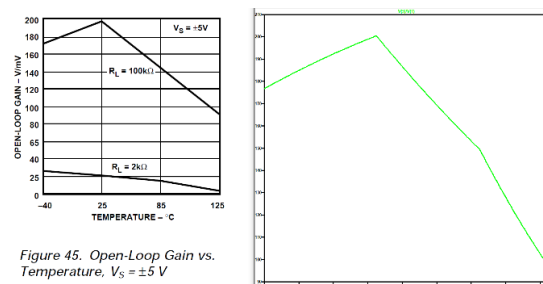


Figure 45. Open-Loop Gain vs. Temperature, $V_S = \pm 5 V$

Рис. 11. Сравнение графика зависимости коэффициента усиления без обратной связи от температуры из документации для ОУ OP191 с графиком, полученным в LTspice

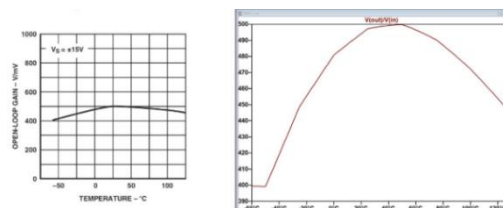


Рис. 12. Сравнение графика зависимости коэффициента усиления без обратной связи от температуры из документации для ОУ LM741 с графиком, полученным в LTspice

При сравнении графиков полученных зависимостей и графиков из описания аналогичных ОУ (рис. 11, 12) можно сделать вывод, что полученные в макромодели с достаточной точностью отражают температурную зависимость коэффициента усиления.

Принцип выявления температурной зависимости у макромодели операционных усилителей с JFET-транзисторами на входе несколько отличается от описанного выше принципа моделирования температурных зависимостей для макромодели на биполярных транзисторах.

В ходе работы было выявлено, что изменения параметров G_a , G_{cm} , G_b , G_c для корректировки влияния температуры на макромодель некорректны. В случае их изменения полностью ломается структура настроенной макромодели и появляются сбои в ее работе.

Решением данной проблемы стало внедрение в схему дополнительного источника питания, который моделирует напряжение смещения на входе в зависимости от температуры (рис. 13). То есть, вносится корректировка параметра в зависимости от температуры, из-за чего вся макромодель становится температурно-зависимой.

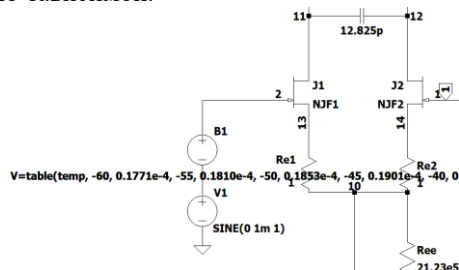


Рис. 13. Дополнительный источник питания B1, моделирующий напряжение смещения входного сигнала в зависимости от температуры

Начальные значения напряжения смещения в зависимости от температуры вносились вручную с большим промежутком исследуемых температур. Далее ряд полученных значений был аппроксимирован в созданной нами программе в среде MatLAB (рис. 10) и получена следующая зависимость напряжения смещения от температуры (рис. 14).

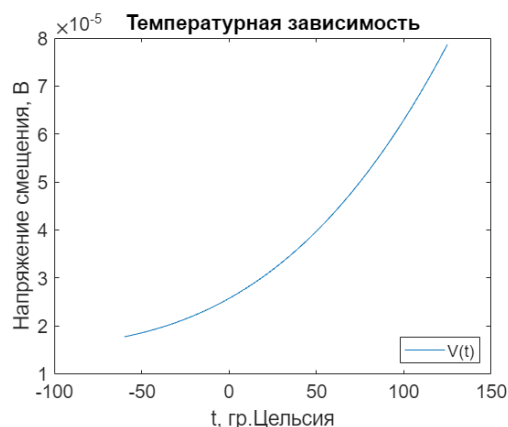


Рис. 14. Зависимость напряжения смещения (напряжения источника В1) от температуры для макромоделей ОУ K544УД2

Также данный метод создания температурной зависимости применим и для макромоделей ОУ на биполярных транзисторах, что и было сделано с макромоделью ОУ K140УД9 (рис 15). Однако, данный метод не является оптимальным в этом случае (может применяться в крайних случаях), так как есть возможность корректировки параметров, отвечающих за температурную зависимость (G_c , G_b),

Применяя данный метод к макромоделям ОУ на биполярных транзисторах создается ряд иных сложностей при дальнейшей работе с ними, поэтому в данной работе такой метод приведен в большей мере для исследовательской деятельности (эксперимента), а не для прикладной.

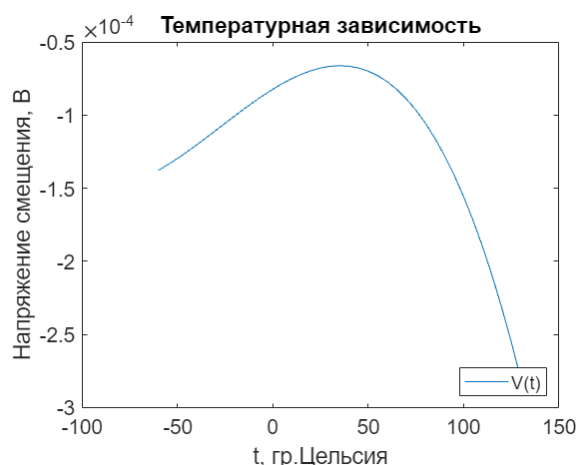


Рис. 15. Зависимость напряжения смещения (напряжения источника В1) от температуры для макромоделей ОУ K140УД9

Промоделированные в LTspice характеристики приведены на рисунках 16, 17.

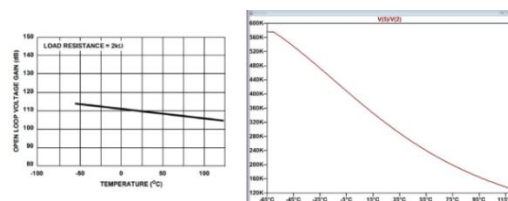


Рис. 16. Сравнение графика зависимости коэффициента усиления без обратной связи от температуры из документации для ОУ K544УД2 с графиком, полученным в Ltspice

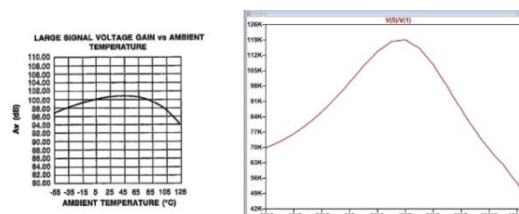


Рис. 17. Сравнение графика зависимости коэффициента усиления без обратной связи от температуры из документации для ОУ K140УД9 с графиком, полученным в Ltspice

Заключение

В ходе работы были получены макромоделей ОУ OP191, LM741, 544УД2, K140УД9 с учетом влияния температуры.

На данный момент в составленных нами макромоделях ОУ OP191, LM741 температурно зависимыми параметрами являются: G_a , G_{cm} , G_b , G_c . В макромоделях ОУ K544УД2, K140УД2 температурно зависимым параметром является дополнительный источник питания.

Сделанные дополнения позволяют учесть температуру в таких параметрах ОУ, как: α_{VD} , V_{OS} , I_{SC} .

Список литературы

1. Micropower Single-Supply Rail-to-Rail Input/Output Op Amps OP191/OP291/OP491 // www.analog.com. - 2006. - Analog Devices
2. Ultralow Offset Voltage Operational Amplifier OP07 // www.analog.com. - 2002. - Analog Devices
3. $\mu A741$ General-Purpose Operational Amplifiers // www.ti.com. - 2018. - Texas Instruments
4. CA3130, CA3130A 15MHz, BiMOS Operational Amplifier with MOSFET Input/CMOS Output // www.renesas.com. - 2005. - Renesas
5. Graeme R. Boyle, Barry M. Cohn, Donald O. Pederson, James E. Solomon. Macromodeling of Operational Integrated Circuit Amplifiers // IEEE Journal Of Solid-State Circuits. - 1974. - VOL.SC 9. - No. 6. - p. 353-364.
6. Фоменкова Е. О. Моделирование интегральных операционных усилителей и исследование их свойств в программе micro-cap // Математическая морфология. Электронный математический и медико-биологический журнал. - 2013. - Том 12. - Вып. 2.
7. Володин В. Я. LTspice: компьютерное моделирование электронных схем. — СПб.: БХВ-Петербург, 2010. — 400 с.

СИСТЕМА РАСПОЗНАВАНИЯ БОРЩЕВИКА СОСНОВСКОГО НА БАЗЕ БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА

Тюрина А.Н., Демокидов А.Р.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе представлены результаты разработки системы для распознавания опасных растений, а именно комплексной защиты сельскохозяйственных земель от Борщевика Сосновского. В основе предложенного метода лежит использование нейронной сети, встроенной в бортовой комплекс беспилотного летательного аппарата для задачи детекции сорняка.

Введение

На сегодняшний день Борщевик Сосновского является одним из самых опасных и быстро распространяющихся сорняков в Российской Федерации. Ежегодно на борьбу с растением в Подмоскowie тратится порядка 500 млн рублей [1], однако положительной динамики достигнуть не удастся. На текущий момент для ликвидации борщевика, в основном, преобладает ручная обработка земель гербицидами. Опасность Борщевика Сосновского заключается в токсичных веществах, содержащихся в соке растений и способных при контакте с кожей вызывать повышенную чувствительность к УФ - излучению и спровоцировать ожоги [2]. Немаловажным следствием бесконтрольного распространения Борщевика является снижение эффективности развития сельскохозяйственной области. Борщевик Сосновского отличается высокой степенью репродуктивной способности. Например, одно растение может производить до 20000 семян [3]. Помимо этого, он устойчив к неблагоприятным условиям климата и может подавлять рост других растений.

Решением данной проблемы является разработка БПЛА (беспилотный летательный аппарат) со встроенной системой распознавания опасного растения с возможностью устранения борщевика без прямого участия человека.

Целью разработки является повышение качества обработки земель от сорняка. В первой главе проводится анализ существующих дронов. Во второй главе описывается структура БПЛА. Третья глава посвящена описанию канала связи. В четвертой главе представляются результаты обучения нейронной сети.

Анализ существующих дронов

Для задачи детекции и последующего устранения борщевика важно, чтобы БПЛА был оснащен камерой и баком с системой распыления. Был проведен анализ рынка современных дронов, применяемых в сельском хозяйстве и выбраны 3 основных кандидата: Joyance JT30L-606, DJI Agras T30, AG116.

Для требуемой задачи был выбран ряд основных параметров дронов: объем бака, время полета, наличие камеры, ширина распыления, наличие GPS модуля и ветроустойчивость. Проведя сравнительный анализ, приведенный в Таблице 1, видно, что лучшей альтернативой является дрон Joyance JT30L-606 [4].

Таблица 1. Анализ дронов

Модель дрона / Параметры	Joyance JT30L-606	DJI Agras T30	AG116
Объем бака, л	30	40	17
Время полета, мин	10-15	8	14
Наличие камеры	Да	Да	Да
Ширина распыления, м	8-10	4-9	7-8
Наличие GPS модуля	Да	Да	Да
Ветроустойчивость, м/с	10	8	-

Структура БПЛА

Структура системы управления дрона представлена на рис. 1. В общем случае, она представлена несколькими блоками: устройство управления, канал связи, полетный контроллер и система FPV (англ. First person view - «от первого лица») [5].

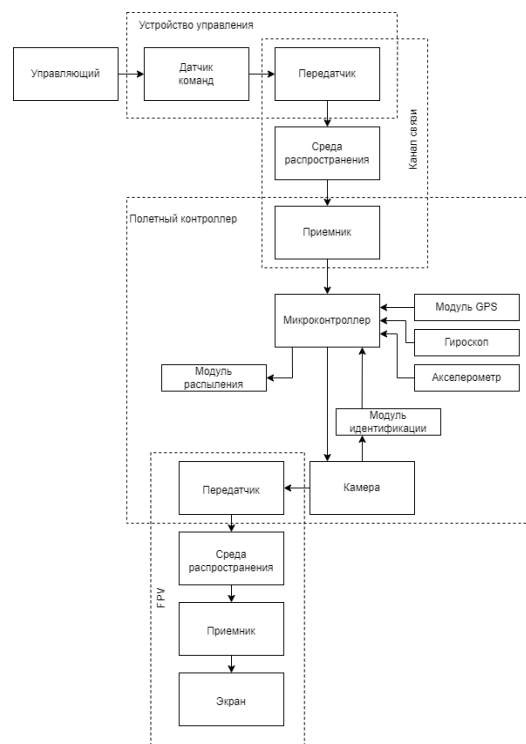


Рис. 1. Блок-схема системы управления дрона

Возможности дрона Joyance JT30L-606 позволяют установить режим работы AV Point Mode, то есть задание маршрута полета БПЛА от точки А до точки В. Далее по каналу связи сигнал передается в блок полетного контроллера, к микроконтроллеру. Модули GPS, гироскоп и акселерометр регулируют полет дрона. В выбранном дроне отсутствует возможность детектировать определенное растение, поэтому дополнительный модуль идентификации, который получает информацию с камеры, добавлен в структуру дрона. Данный модуль связан с модулем распыления, при обнаружении борщевика модуль распыления будет запущен. FPV система дает возможность пользователю с базовой станции наблюдать видеоизображение по дополнительному радиоканалу в режиме реального времени.

Описание канала связи

Для увеличения эффективности передачи информации применяется метод расширения спектра сигнала. В таком случае, информационный сигнал занимает более широкую полосу частот, чем минимально необходимую для передачи данных.

Дрон и базовая станция находятся в системе подвижной радиосвязи, поэтому для увеличения помехозащищенности и обеспечения более надежного соединения применена система с ППРЧ (псевдослучайная перестройка рабочей частоты). ППРЧ характеризуется расширением спектра путем скачкообразного изменения несущей частоты в выделенном для работы системы радиосвязи [6].

Выбор реализации обучения нейронной сети

В настоящее время выделяют три основных задачи машинного обучения: классификация, детекция, сегментация [7]. Для задачи обнаружения борщевика необходимо обучить нейросеть детектировать искомый объект. Наиболее популярной архитектурой нейронной сети для задачи детекции объектов на изображении в режиме реального времени является сверточная сеть YOLO (англ. You only look once - «вы смотрите только один раз») [8]. Из особенностей данной архитектуры можно отметить, что для детекции объектов алгоритму достаточно одно прямое распространение через нейронную сеть. Исходя из этого, прогнозирование всего изображения может выполняться за один запуск алгоритма, в то время как другим нейронным сетям требуется больше.

Из преимуществ данной нейронной сети можно выделить: скорость детекции, что позволяет работать алгоритму в режиме реального времени; высокая точность; повышенный уровень обучаемости, который позволяет ему изучать особенности объектов и применять их при дальнейшей детекции.

Для обучения нейросети был использован набор данных из 250 фотографий борщевика, снятых с высоты полета. Текущий набор данных был собран в районе д. Юрлово Красногорского городского округа Московской области [9]. Обучающая выборка требует наличие специальных меток на каждой фотографии, для этого был использован инструмент labelImg

(см рис. 2). Обучение нейронных сетей проводилось в облачном сервисе Google Colab.



Рис. 2. Разметка фотографий

Результаты обучения нейросети приведены на рис.3.

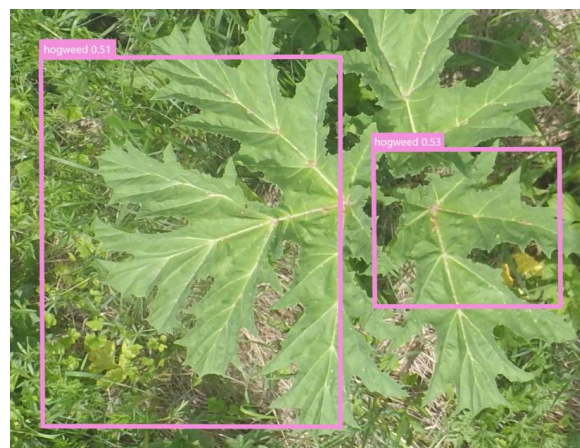


Рис. 3. Результат обучения нейросети YOLO

Заключение

В работе представлены результаты разработки системы для обработки земли от борщевика на базе беспилотного летательного аппарата. В ходе исследования и анализа возможных вариантов, выбран оптимальный дрон и описана структура с дополнительными необходимыми модулями. Дополнительно решена задача разработки нейросети для детектирования борщевика с вероятностью 0.53. Для увеличения вероятности обнаружения будет произведена дополнительная работа.

Список литературы

1. Борьба с борщевиком. ТАСС [Электронный ресурс]. – Режим доступа: <https://tass.ru/moskovskaya-oblast/6258647s>. Дата доступа – 15.01.2023
2. Фомина И. С., Пономаренко Н. В. Борщевик сосновского новый инвазионный вид //Вавиловские чтения" Наследие Н.И. Вавилова в современной науке". – 2019. – С. 152-155.
3. Жиглова О. В., Ходачек О. А. Опыт борьбы с распространением борщевика сосновского на территории Ленинградской области //Региональная экономика и развитие территорий/Под ред. ЛП Совер-

шаевой.–СПб.: ГУАП, 2020, 1 (14).–328 с. – 2020. – С. 190

4. 30 L fumigation drone (JT30L-606) [Электронный ресурс]. – Режим доступа: <https://joyance.tech/ru/9575.html>. Дата доступа – 16.01.2023

5. Джатдоев А. Х. Эволюция систем управления беспилотных летательных аппаратов: от появления до наших дней // Молодой исследователь: вызовы и перспективы. – 2020. – С. 369-374.

6. Борисов В. И., Зинчук В. М., Лимарев А. Е. Помехозащищенность систем радиосвязи с расширением спектра сигналов методом псевдослучайной перестройки рабочей частоты // под ред. В. И. Борисова; изд. 2-е, перераб. и доп. М.: РадиоСофт, 2008. 512 с.

7. Сухов Н. Ю. Решение задач обработки изображений с использованием алгоритмов глубокого обучения. – 2020.

8. Jiang P. et al. A Review of Yolo algorithm developments // Procedia Computer Science. – 2022. – Т. 199. – С. 1066-1073.

9. Набор данных с фотографиями. [Электронный ресурс]. – Режим доступа: <https://github.com/DLopatkin/Heracleum-Dataset>. Дата доступа – 01.12.2022

РАЗРАБОТКА ЦИФРОВЫХ ДВОЙНИКОВ ДЛЯ МОЩНЫХ МОП ТРАНЗИСТОРОВ

Зубкова А.И.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Работа посвящена созданию цифровых двойников (ЦД) для отечественных мощных ДМОП транзисторов 2П829Д и 2П782Ж1. На основе анализа требований к цифровому двойнику, его необходимых частей и результатов измерений реальных образцов МОПТ разработаны их цифровые двойники на языке Python.

Введение

Цифровые двойники важное направление развития во всем мире. Цифровой двойник - это мост между физическим и виртуальным объектом. ЦД позволяют подробно исследовать поведение объектов, не прибегая к использованию реального объекта.

Цифровой двойник включает в себя:

1. Математическая модель объекта;
2. Развивающийся набор данных, относящихся к объекту;
3. Средство динамического обновления или корректировки модели в соответствии с данными. [1]

Разница между цифровым двойником и моделированием в значительной степени зависит от масштаба: моделирование обычно изучает один кон-

кретный процесс, цифровой двойник может сам выполнять любое количество полезных симуляций для изучения нескольких процессов. Цифровой двойник может использоваться на стадиях разработки, производства, эксплуатации устройства или реализации процесса.

Цифровой двойник дает возможность проведения ряда экспериментов без затраченных материалов и средств, с уменьшением времени (так как физические процессы обычно увеличивают время изготовления/эксперимента).

И после проверки на Цифровом двойнике уже можно переходить на физический объект с увеличенной вероятностью правильности действий.

Так как данное направление только начинает развиваться, существует нехватка как специалистов, так и уже созданных платформ или в принципе альтернатив решения поставленных задач. Существующие ЦД не лежат в открытом доступе и проверки работоспособности систем также не существует, что и усложняет задачу, но не уменьшает ее актуальности.

Мощные МОП транзисторы широко используются в схемах управления, преобразователях питания и др. областях. Если для зарубежных компонентов их модели (SPICE, тепловые и др.) можно найти, то для отечественных компонентов эти модели надо создавать [2].

Соответственно, задача состоит в создании ЦД для отечественных мощных ДМОП транзисторов 2П829Д (производитель ОАО «ОКБ «Искра») и 2П782Ж1 (производитель: АО «Воронежский Завод Полупроводниковых Приборов-Сборка»), используемых отечественными производителями электронной аппаратуры.

Анализ литературы показал, что наиболее эффективно реализовать ЦД на языке Python [3]. При этом возможно реализовать эффективное взаимодействие пользователя с ЦД конкретного компонента.

Цифровой двойник транзистора 2П829Д

Для разработки тепловой модели необходимо описать геометрию транзистора, представленную на рис. 1.

Для разработки SPICE модели необходимо измерить характеристики МОПТ и определить параметры его модели

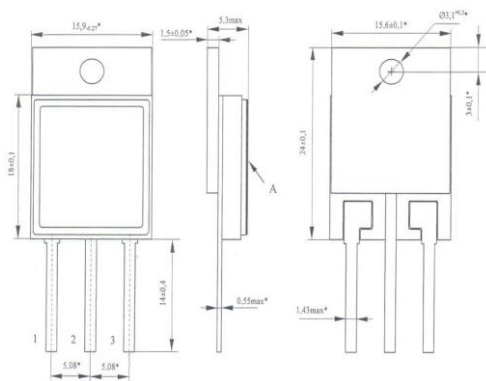


Рис. 1. Геометрия транзистора 2П829Д

Далее были измерены ВАХ транзистора 2П829Д и построен график зависимости. На рисунке 2 изображено ВАХ при различных напряжениях затвор - исток.

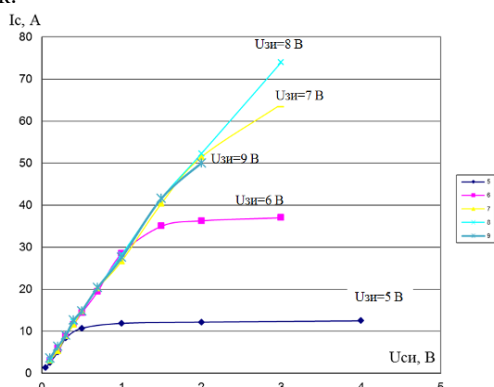


Рис. 2. Измеренная ВАХ транзистора 2П829Д

При сравнении измеренных и смоделированных выходных ВАХ на рисунке 3 можно видеть, что значения максимально сходятся.

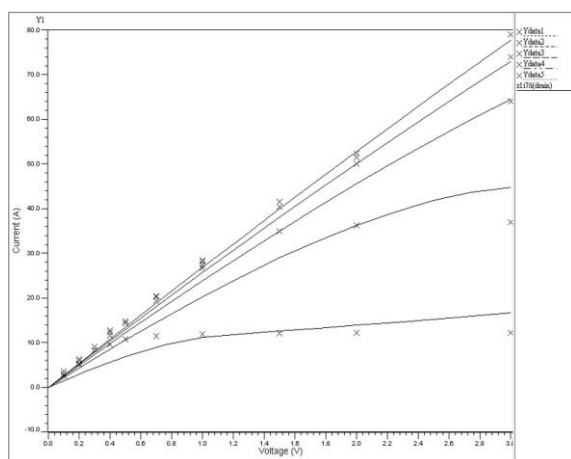


Рис. 3. Сравнение измеренных и смоделированных выходных ВАХ транзистора 2П829Д

Цифровой двойник транзистора 2П782Ж1

Для данного транзистора также можно видеть чертеж его геометрии, представленный на рис. 4.

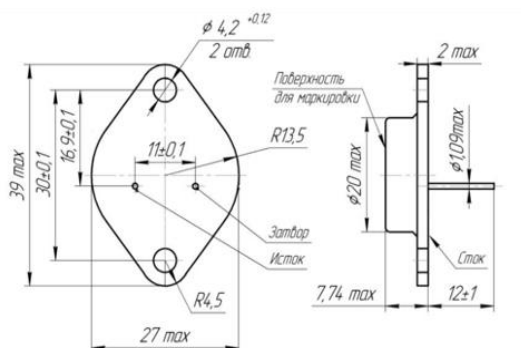


Рис. 4. Геометрия транзистора 2П782Ж1

Для создания его SPICE модели была измерена ВАХ транзистора. На рисунке 5 изображена его ВАХ при различных напряжениях затвор - исток.

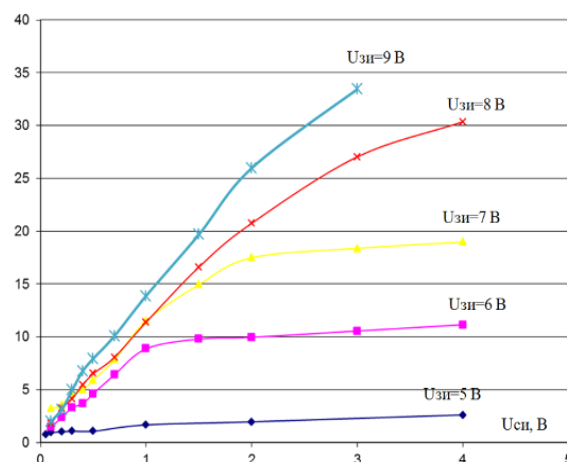


Рис. 5. Измеренная ВАХ транзистора 2П782Ж1

При сравнении измеренных и смоделированных выходных ВАХ на рисунке 6 можно видеть, что значения сходятся менее точно, чем на предыдущем транзисторе, но во всех случаях погрешность мала.

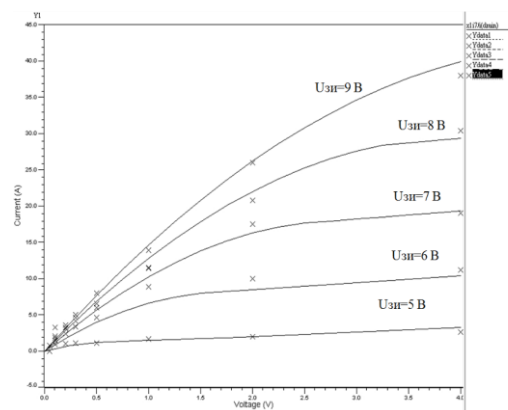


Рис. 6. Сравнение измеренных и смоделированных выходных ВАХ транзистора 2П782Ж1

Заключение

Разработаны части интерфейса и сами цифровые двойники для ряда отечественных мощных ДМОП транзисторов на языке Python.

Предполагается, что результаты работы будут эффективно использованы отечественными производителями электронной аппаратуры.

Список литературы

1. Bailey Brian. The Evolution of Digital Twins - //Semiconductor Engineering, August 3, 2020.
2. Leander Van Cappellen. A Real-time Physics Based Digital Twin for Online MOSFET Condition Monitoring/ Leander Van Cappellen, Martijn Deckers, Omid Alavi - //PV Converter Applications, Thermionic 2022.
3. Tirthajyoti Sarkar. "Digital Twin" with Python: A hands-on example (2021) - URL: <https://towardsdatascience.com/digital-twin-with-python-a-hands-on-example-2a3036124b61> (дата обращения: 25.01.2023).

РАЗРАБОТКА СТРУКТУРЫ ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА ДЛЯ ПРОВЕДЕНИЯ УСКОРЕННЫХ ИСПЫТАНИЙ

Уркунов А.К., Цветков В.Э.,
Ландер Л.Б., Сафина М.Ф.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Данная работа посвящена разработке стенда, который позволит верифицировать основные показатели безотказности ЭРИ. В работе приведено описание структуры стенда и контрольно-измерительного модуля для проведения ускоренных испытаний резисторов, а также построены их структурные схемы. Данное исследование проводится в рамках выполнения проекта по созданию программно-аппаратного комплекса верификации показателей надежности типовых групп резервирования.

Введение

В современных радиоэлектронных средствах (РЭС) для обеспечения надежности количество резервируемых электрорадиоизделий (ЭРИ) в печатных узлах сильно завышено. Этот факт негативно влияет на массогабаритные характеристики устройств. Причиной этому служит недостоверность математических моделей оценки основных показателей надежности вследствие того, что данные интенсивности отказов, полученные в ходе расчетов, не совпадают с результатами, полученными экспериментальным путем [1].

Разработанная в ходе данного исследования структура стенда позволит реализовать впоследствии его настоящий прототип, с помощью которого можно будет проводить ускоренные испытания для валидации уточненной математической модели оценки основных показателей безотказности ЭРИ.

Целью работы является повышение достоверности оценки показателей надежности при проектировании РЭС. Для достижения поставленной цели были изучены теоретические материалы по обеспечению надежности РЭС [2] и мостовым методам измерения сопротивления резисторов, а также построена и проанализирована структура программно-аппаратного комплекса.

Данное исследование состоит из двух частей. В первом разделе описывается структура программно-аппаратного комплекса и указывается предназначение его основных компонентов. Во втором разделе приводится описание структурной схемы контрольно-измерительного модуля и алгоритм расчета сопротивления резисторов.

Структура стенда

Разрабатываемый программно-аппаратный комплекс для проведения ускоренных испытаний резисторов включает в себя 4 основных компонента (рис. 1): печатный узел, климатическая испытательная

камера, контрольно-измерительный модуль и персональный компьютер (ПК).



Рис. 1. Структурная схема стенда

Печатный узел состоит из разъемов и резисторов, являющихся объектом исследования, и мультиплексоров, позволяющих считывать значения напряжения и тока с большого количества резисторов.

Вторым компонентом стенда является климатическая испытательная камера, которая предназначена для контроля надежности ЭРИ. С ее помощью есть возможность варьировать такой параметр среды как температура для проведения ускоренных испытаний. Для данного исследования была выбрана камера Climcontrol M-70/180-80 KTX (рис. 2) [3]. В саму камеру помещен печатный узел, который имеет проводное соединение с контрольно-измерительным модулем, располагающимся вне камеры, через специальное герметичное отверстие.



Рис. 2. Климатическая испытательная камера Climcontrol M-70/180-80 KTX

Контрольно-измерительный модуль предназначен для измерения электрических характеристик ЭРИ, таких как сила тока и напряжение, создания управляющих сигналов и передачи этой информации на ПК. Данный модуль состоит из прецизионных аналоговых ключей, предназначенных для отключения цепи в случае выхода из строя резисторов, диодов Шоттки, используемых для защиты от обратного тока, схемы защиты от короткого замыкания [4], датчиков тока, необходимых непосредственно для измерения значений токов, протекающих через исследуемые резисторы, оптопар, которые выполняют функцию гальванической развязки и предназначены для разделения схемы на силовую и измерительную части. Разрабатываемый контрольно-измерительный модуль предусматривает регулировку трех ускоряющих факторов для проведения испытаний: температуры эксплуатации, коэффициента электрической нагрузки и частоту включения-выключения питания платы с исследуемыми резисторами.

Последним элементом стенда является ПК, который считывает полученную информацию с контрольно-измерительного модуля, обрабатывает ее, а

затем, проанализировав электрические характеристики ЭРИ, делает вывод о том, какие компоненты вышли из строя.

Структура контрольно-измерительного модуля

Одной из ключевых частей разрабатываемого стенда является контрольно-измерительный модуль. В рамках данной научной работы рассматривается проведение ускоренных испытаний в связи с ограниченностью времени, выделяемого на их проведение [5]. Ускорение достигается увеличением нагрузки, подаваемой на объект исследования. Существует множество факторов, изменение которых позволяет сокращать время экспериментов, ускоряя деградиационные процессы в ЭРИ: температура окружающей среды, влажность, механические нагрузки, электрическая нагрузка, частота включений-выключений объектов. В данной научной работе рассматриваются 3 из них:

- температура окружающей среды;
- электрическая нагрузка;
- частота включений-выключений питания экспериментального модуля.

Контрольно-измерительный модуль позволяет управлять данными факторами. Соответственно, для этого необходимы 3 блока:

- блок управления температурой окружающей среды;
- блок управления электрической нагрузкой;
- блок управления частотой включений-выключений питания схемы.

Управление температурой, в частности ее увеличения и уменьшение в процессе проведения испытаний, достигается программным методом. Выбранная климатическая камера Climcontrol M-70/180-80 KTX позволяет изменять температуру в диапазоне от -70°C до $+180^{\circ}\text{C}$. В процессе проведения ускоренных испытаний нижняя граница температуры составляет -55°C , верхняя равна $+125^{\circ}\text{C}$, так как это диапазон рабочих температур исследуемых резисторов [6].

Управление электрической нагрузкой и частотой включений-выключений питания платы с резисторами реализовано на программном уровне. Контрольно-измерительный модуль позволяет увеличить коэффициент электрической нагрузки до 1 и подавать питание с частотой 1 Гц.

Изменение данных трех факторов позволяет достичь необходимого для определения наработки на отказ рассматриваемых резисторов согласно отечественному справочнику по надежности 2006 года [7].

Кроме того, для проведения эксперимента необходимо не только управлять изменением выбранных параметров, но и измерять некоторые величины. Основной характеристикой резисторов является электрическое сопротивление. Изначально было решено реализовать схему измерения сопротивления на основе одного из методов [8]:

- Мостовая схема Кельвина;
- Мостовая схема Уитстона;
- Мостовой метод Меггера.

Каждый из методов подходит для измерения сопротивления постоянного тока в различных диапазо-

нах: низком (до 1 Ом), среднем (от 1 Ом до 100 кОм) и высоком (более 100 кОм). Однако для оптимизации контрольно-измерительного модуля был разработан алгоритм, позволяющий рассчитывать сопротивление резисторов на основе значений силы тока и электрического напряжения и реализованный на программном уровне. На рисунке 3 представлена структура контрольно-измерительного модуля:

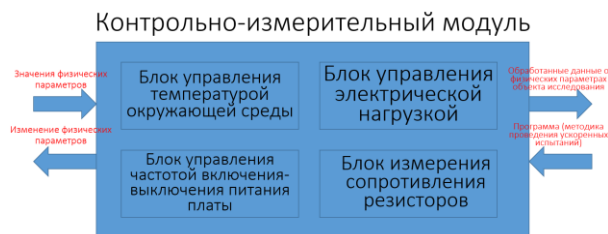


Рис. 3. Структурная схема контрольно-измерительного модуля

Таким образом, контрольно-измерительный модуль состоит из четырех основных блоков, позволяющих управлять выбранными параметрами, увеличивая или уменьшая их значения, и блока измерения сопротивления, позволяющим определять значение данной величины на исследуемых резисторах.

Заключение

В данной работе была проведена разработка и исследование структуры программно-аппаратного комплекса для проведения ускоренных испытаний, в частности самого стенда и контрольно-измерительного модуля.

Таким образом, верификация математических моделей оценки показателей надёжности позволяет сократить количество ЭРИ, используемых в РЭС и тем самым сделать устройства более компактными и легкими, соответствующими современным требованиям надежности электроники.

Список литературы

1. Кулибаба, А. Я. Анализ влияния входного контроля и дополнительных испытаний на надежность электронной компонентной базы / А. Я. Кулибаба, А. А. Сашов, М. К. Суконкин, А. Ю. Штукарев. - Текст: электронный // Ракетно-космическое приборостроение и информационные системы. - 2020. - № 3 (7). - С. 87-92. - URL: https://spacedevice.ru/wp-content/uploads/2020/09/10_p87_0703.pdf (дата обращения 15.01.2023).
2. RIAC-HDBK-217Plus. Handbook of 217PlusTM reliability prediction models. USA: RIAC, 2006. 170 p.
3. Испытательная Климатическая камера Climcontrol M-70/180-80 KTX. - Текст: электронный // Мир оборудования: сайт. - 2022. - URL: <https://mirobordovaniya.ru/product/ispytatelnaya-klimaticheskaya-kamera-climcontrol-m-70-180-80-kth/> (дата обращения: 10.01.2023).
4. Становой, А. Схемы защиты от короткого замыкания и перегрузок в блоке питания / А. Становой. - Текст: электронный // Запитка: [сайт]. - 2022. - 10 февр. - URL: <https://zapitka.ru/pitanie/shema->

zaschity-ot-korotkogo-zamykaniya-v-bloke-pitaniya
(дата обращения: 11.01.2023).

5. РД 50-424-83. Надежность в технике. Ускоренные испытания. Основные положения: руководящий нормативный документ: издание официальное: утвержден и введен в действие Постановлением Государственного комитета СССР по стандартам от 10 октября 1983 г. № 4903: введен впервые: дата введения 1985-01-01 / разработан Государственным комитетом СССР по стандартам. - Москва: Издательство стандартов, 1985. - 12 с. - Текст: непосредственный.

6. CF-25 (C1-4) 0.25 Вт, 36 Ом, 5%, Резистор углеродистый. - Текст: электронный // ЧИП и ДИП: сайт. - 2023. - URL: <https://www.chipdip.ru/product0/41290> (дата обращения: 10.01.2023).

7. Надежность ЭРИ 2006: справочник. Москва: МО РФ, 2006. С. 641.

8. Measurement of Resistance. - Текст: электронный // Electrical4U: сайт. - 2020. - URL: <https://www.electrical4u.com/measurement-of-resistance> (дата обращения: 11.01.2022).

РАЗРАБОТКА МЕТОДИКИ УТОЧНЕННОЙ ОЦЕНКИ ПОКАЗАТЕЛЕЙ БЕЗОТКАЗНОСТИ СОВРЕМЕННЫХ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

Демокидов А.Р., Тюрина А.Н.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В научно-исследовательской работе решена первоочередная научная проблема в области обеспечения и оценки единичных показателей надежности электронных модулей первого уровня (содержащихся в космических системах), связанная с недостаточностью численной оценки показателей безотказности современных интегральных микросхем, входящих в их состав. Для решения указанной научной проблемы проведен обзор и анализ статистики отказов интегральных микросхем и методик оценки показателей безотказности; предложена новая методика уточненной оценки показателей безотказности современных интегральных микросхем; проведена апробация разработанной методики для конкретного объекта исследования.

Введение

На сегодняшний день ключевую роль в развитии экономики Российской Федерации играет совершенствование ракетно-космической отрасли [1]. Перед страной стоит задача обеспечения конкурентоспособности и рентабельности космических услуг связи [2].

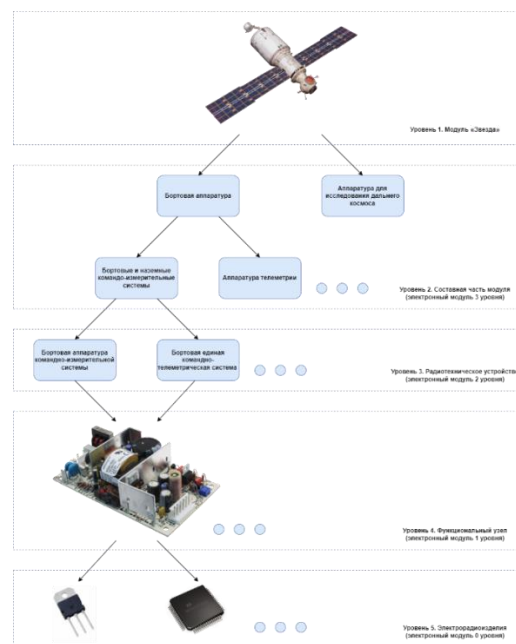


Рис. 1. Структура космического модуля «Звезда»

Для работы космических систем необходимо соблюдение требований к надежности и безопасности [3]. Однако существует множество проблем с непредвиденными отказами космических систем во время их эксплуатации по причине возникновения отказов на самом низшем уровне (рисунок 1) – электрорадиоизделий [4].

Аналитическим решением указанных проблем занимались и занимаются такие известные люди, как: И.А. Ушаков, Б.А. Гнеденко, Б.В. Козлов, Г.В. Дружинин, а также другие советские и российские ученые. Однако в методах оценки параметров безотказности по ходу настоящего исследования выявлены недостатки, а именно отсутствие единых формул расчета интенсивности отказов кристалла для интегральных микросхем, эксплуатационной интенсивности отказов.

Поэтому актуальным становится исследование математических моделей оценки показателей безотказности современных интегральных микросхем (ИМС) и разработка новой методики оценки показателей безотказности.

Целью исследования является повышение точности оценки показателей безотказности ИМС. Актуальность состоит в построении методики уточненной оценки показателей безотказности, учитывающей все необходимые актуальные параметры из отчетов по надежности фирм-производителей, исходя из реально проведенных испытаний.

Объектами исследования являются математические модели параметров безотказности ИМС, а именно коэффициент ускорения A , интенсивность отказов кристалла $\lambda_{кр}$, температура кристалла T , температурный коэффициент K , интенсивность отказов корпуса $\lambda_{корп}$ и эксплуатационная интенсивность отказов λ_3 .

Исследование статистики отказов ЭРИ

Число отказов каждого устройства (электронный модуль 2 уровня) производится суммированием числа отказов каждой платы, входящей в данное устройство, и числа отказов электрорадиоизделий (ЭРИ), входящих в состав данного устройства [5]. Отказы интегральных микросхем, связанные с утечками, пробоями и перегревами, являются одним из наиболее распространенных видов отказов современного радиоэлектронного средства [6]. Именно поэтому необходимо более точно рассчитывать параметры безотказности ИМС.

Центр исследования надежности RIAC [7] опубликовал статистику о процентном соотношении категорий отказов ЭРИ (рисунок 2).

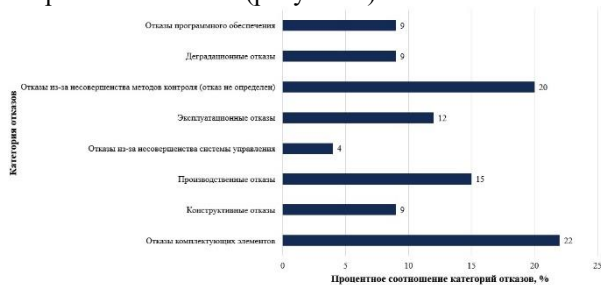


Рис. 2. Распределение категорий отказов

Проанализировав данные из рисунков, можно обнаружить, что у 20% отказов причины не установлены [8]. Отсюда можно сделать вывод, что следствием отказов служит недостоверный и неточный расчет параметров безотказности ЭРИ.

Описание методики оценки параметров безотказности ИМС

Разработанная методика представлена в виде иерархической диаграммы (рисунок 3).

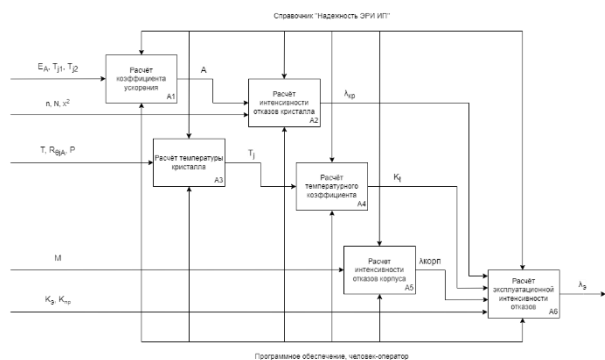


Рис. 3. Методика оценки параметров безотказности

Разработанная методика используется для расчета показателей безотказности ИМС. Для формирования исходных данных для блоков необходимо воспользоваться отчетом надежности производителя ЭРИ и технической документацией. Все формулы для расчета берутся из справочника «Надежность ЭРИ ИП» [9].

Блок A1 – Коэффициент ускорения определяется

по уравнению Аррениуса: $A = e^{\left(\frac{E_a}{k} \left(\frac{1}{T_{j1}} - \frac{1}{T_{j2}}\right)\right)}$, где E_a – энергия активации, k – постоянная Больцмана, T_{j1} – рабочая температура, T_{j2} – предельная температура.

Блок A2 – Интенсивность отказов кристалла вычисляется по формуле: $\lambda_{кр} = \frac{\chi^2}{2 \cdot n \cdot N \cdot A}$, где χ^2 – значение хи-квадрат распределения при заданной доверительной вероятности с $(2 \cdot f + 2)$ степенями свободы – количества отказов, n – размер выборки, N – суммарное количество часов испытаний.

Блок A3 – Температура кристалла определяется из соотношения: $T_j = T + R_{\theta ja} \cdot P$, где T – температура окружающей среды, $R_{\theta ja}$ – тепловое сопротивление «кристалл - окружающая среда», P – фактическая мощность рассеяния ИМС.

Блок A4 – Температурный коэффициент вычисляется по формуле: $K_t = 0.1 \cdot e^{-\frac{E_a}{k} \left(\frac{1}{T_j + 273} - \frac{1}{298}\right)}$, где E_a – энергия активации, k – постоянная Больцмана, T_j – температура кристалла.

Блок A5 – Интенсивность отказов корпуса определяется из соотношения: $\lambda_{корп} = 2.8 \cdot 10^{-10} \cdot M^{1.08}$, где M – количество выводов ИМС.

Блок A6 – Эксплуатационная интенсивность отказов вычисляется по формуле: $\lambda_3 = (\lambda_{кр} \cdot K_t + \lambda_{корп} \cdot K_3) \cdot K_{np}$, где $\lambda_{кр}$ – интенсивность отказов кристалла, K_t – температурный коэффициент, $\lambda_{корп}$ – интенсивность отказов корпуса, K_3 – коэффициент эксплуатации, K_{np} – коэффициент качества.

Выбор ЭРИ для исследования и обоснование выбора математического аппарата

Для расчета рассмотрена ИМС SN74AHC1G00BVR производителя Texas Instruments [10]. Заполним таблицу 1 согласно информации из отчета производителя и справочника «Надежность ЭРИ ИП» [9].

Таблица 1. Данные для расчета

Параметр	Значение	Ед. измерения
T_{j1}	328	K
T_{j2}	398	K
n	10859	шт
N	1000	ч
P	0.04	Bm
$R_{\theta ja}$	240	°C/Bm
M	5	шт
K_3	0.5	
K_{np}	1	

Наработка на отказ, полученная в результате испытаний $MTBF = 9.3 \cdot 10^8$ ч.

Используя разработанную методику для вычисления параметров безотказности выбранной ИМС и данные из таблицы 1 получим следующие значения, которые запишем в таблицу 2.

Таблица 2. Результаты расчета

Параметр	Значение	Ед. измерения
A	77.935	
$\lambda_{кр}$	$1.08 \cdot 10^{-9}$	1/ч
T_i	34.6	°C
K_t	0.234	
$\lambda_{корп}$	$1.592 \cdot 10^{-9}$	1/ч
λ_3	$1.049 \cdot 10^{-9}$	1/ч

Таким образом, наработка на отказ $MTBF$ при расчете по разработанной методике

$$MTBF = \frac{1}{\lambda_3} = 9.53 \cdot 10^8 \text{ ч.}$$

Рассчитаем наработку на отказ $MTBF$ используя справочники «Надёжность электрорадиоизделий 2006» [11], MIL-HDBK-217F [12] и рассчитаем относительную погрешность (таблица 3).

Таблица 3. Расчет наработки на отказ

Справочники	MTBF	δ_{MTBF}
Надёжность ЭРИ ИП	$9.49 \cdot 10^8$ ч	2.04%
Надёжность ЭРИ 2006	$0.24 \cdot 10^8$ ч	97.4%
MIL-HDBK	$9.99 \cdot 10^5$ ч	99.89%

Так как относительная погрешность вычисления эксплуатационной интенсивности отказов по справочнику «Надёжность ЭРИ ИП» много меньше погрешности при использовании формулы из справочника «Надёжность электрорадиоизделий 2006» и «MIL-HDBK», для вычисления эксплуатационной интенсивности отказов необходимо пользоваться моделью из справочника «Надёжность ЭРИ ИП».

Заключение

В результате проведённого исследования была разработана методика уточненной расчетной оценки показателей безотказности современных ИМС. В работе приведен расчет для микросхемы SN74AHC1G00, чтобы подтвердить корректность предложенной методики. В процессе выполнения решены следующие задачи: уточнена оптимальная модель расчета интенсивности отказов кристалла; уточнена наиболее достоверная модель расчета эксплуатационной интенсивности отказов; проведено тестирование разработанной методики.

Достоинства методики: высокая точность оценки параметров безотказности (для выбранного объекта исследования относительная погрешность составляет 2%).

Недостатки методики: расчет параметров вручную.

Для устранения главного недостатка ведется работа по разработке программного обеспечения для автоматизации расчета параметров.

Список литературы

1. Основные положения Федеральной космической программы 2016 – 2025. [Электронный ресурс]. URL: <https://www.roscosmos.ru/22347/> (Дата обращения 05.01.2023).
2. Основные положения государственной политики Российской Федерации в области космической деятельности на период до 2030 года и дальнейшую перспективу. [Электронный ресурс]: URL: <https://legalacts.ru/doc/osnovnye-polozenija-osnov-gosudarstvennoi-politiki-rossiiskoi-federatsii/> (Дата обращения 05.01.2023).
3. ГОСТ Р 56526-2015. Требования надежности и безопасности космических систем, комплексов и автоматических космических аппаратов единичного (мелкосерийного) изготовления с длительными сроками активного существования. - М.: Стандартинформ, 2016. – 55 с.
4. Королев П. С. Комплексный метод оценки показателей безотказности радиотехнических устройств космической аппаратуры //Известия высших учебных заведений. Приборостроение. – 2021. – Т. 64. – №. 4. – С. 316-328.
5. Севастьянов Н.Н., Андреев А.И. Основы управления надежностью космических аппаратов с длительными сроками эксплуатации / под общ. ред. Н.Н. Севастьянова. – Томск: Издательский Дом ТГУ, 2015. – 266 с.
6. Андреев А.И., Жаднов В.В., Кофанов Ю.Н. Виды и причины отказов РЭС: учеб. пособие. М.: МГИЭМ, 1995. 85 с.
7. RIAC-HDBK-217Plus. Handbook of 217PlusTM reliability prediction models. USA: RIAC, 2006. 170 p.
8. Королев П. С. Влияние коэффициента качества производства аппаратуры на оценку показателей надежности радиолокационного оборудования в RIAC-HDBK-217PLUS // В кн.: Инновационные, информационные и коммуникационные технологии: сборник трудов XVI Международной научно-практической конференции / Отв. ред.: И. А. Иванов. Ассоциация выпускников и сотрудников ВВИА им. проф. Жуковского, 2019. С. 439441.
9. Надёжность ЭРИ ИП: справочник. – М.:МО RF, p. 52, 2006. 25
10. SN74AHC1G00 Quality, reliability [Электронный ресурс]. URL: <https://www.ti.com/quality-reliability-packaging-download/report?opn=SN74AHC1G00DBVR> (Дата обращения 05.01.2023).
11. Надёжность ЭРИ: справочник. – М.:МО RF, p. 641, 2006.
12. Military Handbook (1991). Reliability prediction of electronic equipment. MIL-HDBK-217F, 205 p.

МЕТОД ПЕРЕДАЧИ ИНФОРМАЦИИ В УСЛОВИЯХ ВОЗДЕЙСТВИЯ МОЩНЫХ ШИРОКОПОЛОСНЫХ ПОМЕХ

Артамонов Д.О., Уткин Б.В.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Существуют различные методы подавления радиопомех в средствах передачи информации. В частности, помех, превышающих уровень полезного сигнала. В данной работе приведены результаты исследования метода подавления широкополосных помех с выделением узкополосного полезного сигнала. Результаты исследования показывают высокую эффективность метода, который позволит значительно улучшить помехозащищенность систем передачи информации.

Введение

В радиотехнике существует ряд важных параметров, оказывающих основное влияние на работоспособность системы, а именно диапазон частот, в котором передается информация, помехоустойчивость системы, дальность действия. Другие параметры системы обычно определяются ее целевым назначением.

В рамках данной работы ставится задача на достижение максимальной помехозащищенности системы при использовании минимальной полосы пропускания канала, которая определяет диапазон частот, используемый для передачи информации.

В теореме Шеннона-Хартли устанавливается ограничение на максимальный объем безошибочных цифровых данных, которые могут быть переданы по такому каналу [1].

Значение спектральной эффективности канала в соответствии с теоремой определяется как:

$$C = B \cdot \log_2 \left(1 + \frac{S}{N} \right) \quad (1)$$

где C - пропускная способность канала, B - полоса пропускания канала, S - мощность сигнала и N - мощность шума.

Для узкополосного сигнала, где спектральная эффективность передачи информации, например, составляет 0,25 бит/с/Гц, отношение сигнал/шум составляет минус 7,2 дБ. По современным стандартам этого результата недостаточно для обеспечения высокой помехозащищенности системы.

Известны способы, позволяющие достичь производительности системы с отношением сигнал / (помеха + шум) в диапазоне от минус 7,2 до минус 20 дБ. Например, способ подавления помех с использованием адаптивного пространственного фильтра на основе антенной решетки [2,3] или методы согласованной и адаптивной нелинейной фильтрации [3,4]. Или методы, основанные на широкополосных сигналах [5].

Однако в большинстве случаев известные методы направлены на борьбу с помехами, чьи параметры

заведомо известны, например, частотная область воздействия, время возникновения, пространственное направление, поляризация и т.д. И на противодействие помехам, параметры которых неизвестны, но при этом жертвуя одним из важнейших ресурсов радиоустройства – эффективным использованием радиочастотного диапазона.

Из вышеперечисленных соображений вытекает необходимость проводить исследования в области многоканальной обработки с использованием узкополосных сигналов.

Для решения этой задачи предлагается алгоритм цифровой обработки сигналов, основанный на методе двухканального синхронного приема. Согласно этому алгоритму, проводится выделение узкополосного полезного сигнала с подавлением воздействующей помехи.

Математическая модель метода

Современная радиоэлектронная элементная база, а также вычислительные мощности систем на кристалле позволяют реализовать двухканальный синхронный приемник.

На рисунке 1 показана модель формирования сигналов на входах двухканального приемника.

Алгоритм цифровой обработки сигнала сводится к следующим этапам:

- Вычисление значений коэффициентов α_1 и α_2 , которые получены по опорным сигналам. Эти коэффициенты обеспечивают максимальное значение отношения сигнал / шум.

- Оценка информационных сигналов на основе полученных коэффициентов, когда они стационарны на некотором временном интервале.

В свою очередь оценка сигналов сводится к следующим пунктам:

- Вычисление координат вектора оценки информационного сигнала:

$$\vec{S} = \alpha_1 \cdot \vec{f}_1 + \alpha_2 \cdot \vec{f}_2 \quad (2)$$

- В области всех информационных сигналов найти такой, который имеет наименьшее расстояние до вычисленной:

$$\vec{S}_k: d_k = \min_{i=1..N} d(\vec{S}_i, \vec{S}) \quad (3)$$

Значения коэффициентов для достижения оптимального отношения сигнал / шум определяются по формулам (4) и (5):

$$\alpha_1(\vec{f}_1, \vec{f}_2, \vec{S}^*) = \frac{|\vec{f}_2|^2(\vec{f}_1, \vec{S}^*) - (\vec{f}_2, \vec{S}^*)(\vec{f}_1, \vec{f}_2)}{|\vec{f}_1|^2|\vec{f}_2|^2 - (\vec{f}_1, \vec{f}_2)(\vec{f}_1, \vec{f}_2^*)} \quad (4)$$

$$\alpha_2(\vec{f}_1, \vec{f}_2, \vec{S}^*) = \frac{|\vec{f}_1|^2(\vec{f}_2, \vec{S}^*) - (\vec{f}_1, \vec{S}^*)(\vec{f}_1, \vec{f}_2)}{|\vec{f}_1|^2|\vec{f}_2|^2 - (\vec{f}_1, \vec{f}_2)(\vec{f}_1, \vec{f}_2^*)} \quad (5)$$

где \vec{f}_1 и \vec{f}_2 – принятые сигналы, \vec{S}^* - вектор, комплексно-сопряженный с вектором переданного сигнала \vec{S} , (\vec{f}_i, \vec{S}^*) – скалярное произведение n -мерных комплексных векторов, а $\alpha_1(\vec{f}_1, \vec{f}_2, \vec{S}^*)$ и $\alpha_2(\vec{f}_1, \vec{f}_2, \vec{S}^*)$ - комплексные функции от соответствующих аргументов.

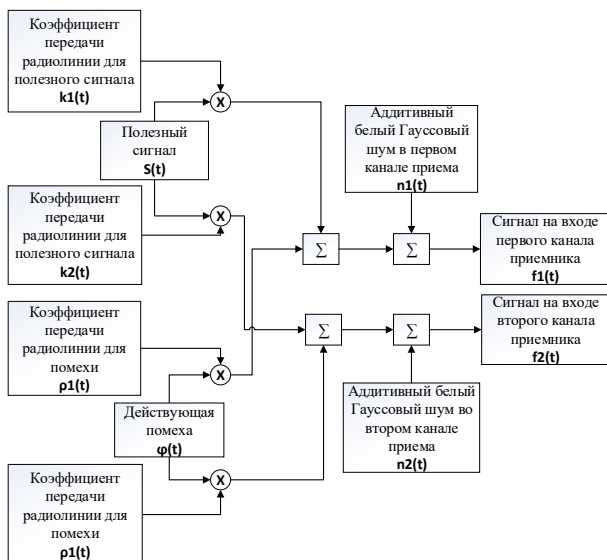


Рис. 1. Модель формирования сигналов на входах приемника

Программная модель метода

Программная модель реализована на языке программирования C++ и состоит из программных функциональных блоков, выполняющих цифровую обработку сигналов.

Модель реализует обработку сигнала на нулевой промежуточной частоте. Сигнал представляет собой выборку, полученную на частоте дискретизации 1 МГц без учета ошибок квантования.

Моделируемая система передачи информации, изображенная на рисунке 2, включает в себя следующие блоки:

- Передатчик
- Приемник
- Имитатор помехового сигнала
- Радиоканал

Имитатор помехового сигнала генерирует сигналы помехи и шума в виде белого Гауссова шума. Сигнал помехи в отличие от шума является источником точечного когерентного сигнала, который коррелируется на входах приемника. Он реализован в виде набора случайных чисел Гаусса, который дает равномерный спектр и независимые выборки сигнала.

Радиоканал моделирует распространение сигнала с учетом эффектов Доплера и многолучевости.

Модель приемника осуществляет обработку сигнала в соответствии с исследуемым методом, вычисление отношения сигнал / (помеха + шум) и блочной ошибки передачи информации.

В структуре полезного сигнала содержатся опорные символы для вычисления необходимых коэффициентов α_1 и α_2 .

Отношение сигнал / (помеха + шум) для исследуемой модели рассчитывается в соответствии с выражением:

$$SNR = \frac{P_S \cdot |\alpha_1 k_1 + \alpha_2 k_2|^2}{P_\phi \cdot |\alpha_1 \rho_1 + \alpha_2 \rho_2|^2 + P_{n1} \cdot |\alpha_1|^2 + P_{n2} \cdot |\alpha_2|^2} \quad (6)$$

где P_S и P_ϕ – мощности полезного сигнала и помехи, нормированные на единую мощность, P_{n_i} – мощности шума в первом и втором каналах приемника, нормированные на единую мощность, k_i и ρ_i – соответствующие коэффициенты передачи для полезного сигнала и помехи в канале связи.

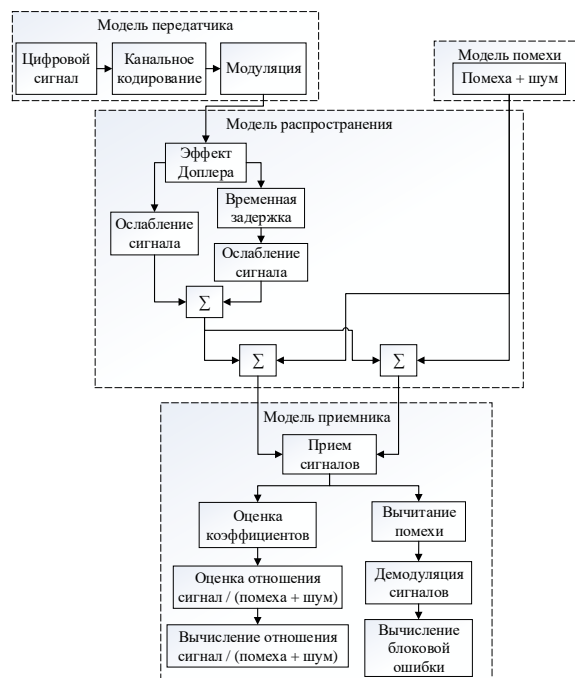


Рис. 2. Блок-схема программной модели

На рисунке 3 приведены результаты моделирования системы в виде зависимости вероятности блочной ошибки P в принятых данных при различном количестве опорных символов N в составе полезного сигнала для двух значений SNR .

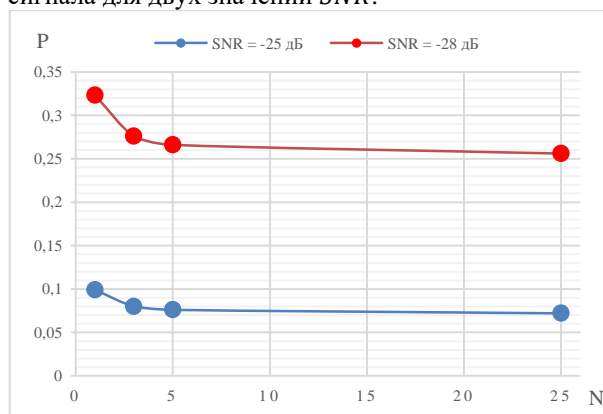


Рис. 3. Результаты моделирования системы

Полученные результаты демонстрируют высокую эффективность предлагаемого метода двухканального синхронного приема с выделением полезного сигнала при отношении сигнал / помеха более минус 20 дБ.

Эксперименты с увеличением количества опорных символов показывают незначительность их влияния при количестве больше пяти.

Заключение

Постоянное увеличение количества радиотехнических средств и методов передачи информации неизбежно приводит к усложнению радиоэлектронной обстановки и повышает актуальность проблемы помехоустойчивости.

Существующие методы фильтрации и выделения полезного сигнала ориентированы на работу с помехами, чьи параметры известны, либо на работу с помехами, чьи параметры неизвестны, но жертвуя эффективностью использования радиочастотного диапазона.

Полученные результаты программной модели показывают возможность передачи узкополосного сигнала при воздействии помех с широкой полосой спектра и превышением над уровнем полезного сигнала до 28 дБ, что демонстрирует его существенное преимущество по сравнению с другими методами.

Список литературы

1. Скляр, Бернард. Цифровая связь. Теоретические основы и практическое применение. – М.: Издательский дом «Вильямс», 2003. – 1104 с.
2. Ермолаев В.Т., Семенов В.Ю., Флакман А.Г., Ястребов А.В. Пространственно-временной компенсатор широкополосных помех на основе метода степенных векторов. Известия высших учебных заведений России. Радиоэлектроника. 2017;(4):19-26.
3. Джиган В. И. Адаптивная фильтрация сигналов: теория и алгоритмы. – М.: Техносфера, 2013. 528 с.
4. Милащенко Е.А., Язовский А.А. Метод адаптивного нелинейного подавления негауссовских помех в когерентном полосовом тракте приемника. Труды учебных заведений связи. 2019;5(4):49-57.
5. Дмитриев А.С., Мохсени Т.И., Сьерра-Теран К.М. Сверх- и гиперширокополосная относительная передача информации на основе хаотических радиоимпульсов // Известия вузов. Прикладная нелинейная динамика. 2018. Т. 26, № 4. С. 59–74.

СОЗДАНИЕ ПРОГРАММЫ УЧЕТА ПРИ СХЕМОТЕХНИЧЕСКОМ МОДЕЛИРОВАНИИ ВОЗДЕЙСТВИЯ НА РАБОТУ КМОП СХЕМ ВНУТРЕННИХ ФАКТОРОВ СТАРЕНИЯ

Кобяков В.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Данная работа посвящена созданию программы для автоматизации возможности учета при схемотехническом моделировании воздействия на работу КМОП схем внутренних факторов старения.

Приведены результаты использования разработанной программы для анализа эффекта горячих но-

сителей для схемы КМОП инвертора с проектными нормами 30 нм.

Введение

С течением времени работы характеристики и параметры КМОП транзистора ухудшаются из-за факторов старения [1]. В литературе имеются выражения для учета влияния старения на МОП транзисторы [2].

Но современные средства схемотехнического моделирования не учитывают эти эффекты и не проводят оценку их влияния на работу схем.

Для автоматизации анализа процессов старения КМОП транзисторов на этапе схемотехнического моделирования схемы возникает необходимость в разработке вспомогательной программы к пакету Spice [3].

Главной целью работы является создание программы для автоматизации учета при схемотехническом моделировании воздействия на работу КМОП схем внутренних факторов старения.

Для этого решались следующие задачи:

- обзор основных эффектов старения КМОП транзисторов;
- разработка вспомогательной к LtSpice программы, автоматизирующей оценку и учет старения транзисторов;
- добавление возможности учета температуры в процессе оценки старения;
- создание руководства пользователя по работе с программой.

Основные эффекты старения

Основными эффектами старения КМОП транзисторов являются:

- 1) Инжекция горячих носителей (англ. hot-carrier injection)
- 2) Температурная нестабильность с отрицательным смещением (англ. negative-bias temperature instability) [4]

НСИ эффект проявляется в моменты скачков тока стока и тока подложки при переключении.

Эффекты NBTI и PBTI проявляются при обратных напряжениях на затворах МОПТ.

Поэтому для анализа степени проявления эффектов старения при работе схемы необходимо скачки тока и напряжения проинтегрировать за время моделирования.

Функции вспомогательной программы следующие:

- 1) Подготовить команды .measure, которые интегрируют скачки токов для транзисторов схемы
- 2) Проанализировать после проведения моделирования полученные интегралы
- 3) Определить степень старения транзисторов по полученным интегралам.

Работа программы

Программа реализована на языке Python.

Иллюстрация работы программы для схемы инвертора в LtSpice приведена ниже (см. рис. 1).

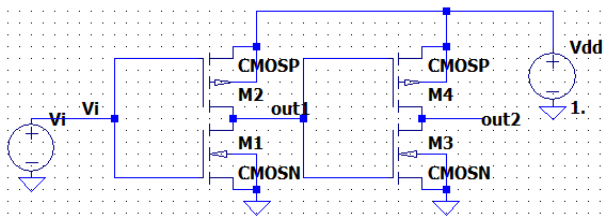


Рис. 1. Схема инвертора

Для анализа старения транзистора M1 необходимы скачки тока стока I_d и тока подложки $I(B1ib_)$ при напряжениях на входе $V(vi)$ и на выходе $V(out1)$ транзистора (см. рис. 2).

Пример интерфейса программы приведен ниже (см. рис. 3).

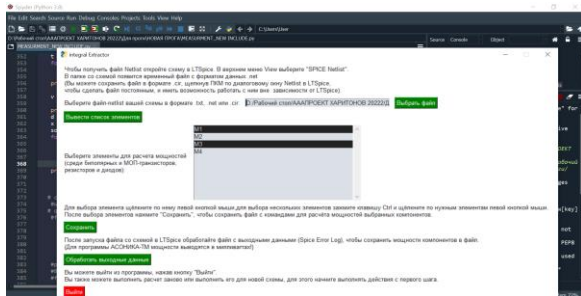


Рис. 2. Интерфейс программы

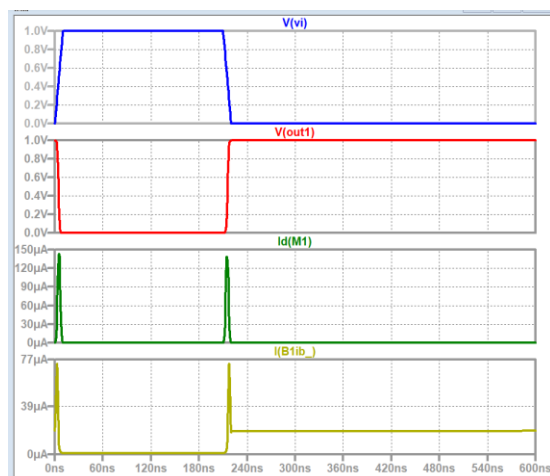


Рис. 3. Графики $V(vi)$, $V(out1)$, $I_d(M1)$, $I(B1ib_)$ из LtSpice

Для интегрирования скачков токов используются команды .measure (см. рис. 4), которые автоматически генерируются вспомогательной программой.

```
.measure tran HCI_M1 integ I(B1ib)
.measure tran maxHCI_M1 max I(B1ib)
.measure tran maxId_M1 max Id(M1)
```

Рис. 4. Сгенерированные команды для расчета интегрирования скачков токов в программе LtSpice

Из необходимости составления include-файла возникает потребность в автоматизации процесса.

По результатам симуляции LtSpice программа анализирует степень проявления эффекта горячих носителей для каждого транзистора и показывает эти данные пользователю в порядке убывания их проявления.

Заключение

На данный момент разработанная программа автоматизирует составление include файлов для вывода интегралов скачков тока, вызванных горячими носителями, и анализирует полученный результат.

В перспективе и в процессе разработки находятся включение во вспомогательную программу:

- 1) Учета эффектов NBTI, PBTI
- 2) Учета температуры при моделировании схемы
- 3) Корректировка параметров транзисторов с учетом старения

Список литературы

1. J.-M. Baek, Y.-C. Kim, and M.-T. Cho, "Hot-Carrier Degradation of NMOSFET," Journal of the Korea Academia-Industrial cooperation Society, vol. 10, no. 12, pp. 3626–3631, Dec. 2009, doi: 10.5762/kais.2009.10.12.3626.
2. B. S. Doyle and K. R. Mistry, "A general gate-current p-MOS lifetime prediction method applicable to different channel structures," IEEE Electron Device Letters, vol. 11, no. 11, pp. 547–548, Nov. 1990, doi: 10.1109/55.63028.
3. P. Ptak, "Application of the software package ltspace for designing and analysing the operation of electronic systems," society. integration. education. Proceedings of the International Scientific Conference, vol. 5, pp. 402–409, May 2018, doi: 10.17770/sie2018vol1.3118.
4. A. Teramoto, R. Kuroda, and T. Ohmi, "NBTI Mechanism Based on Hole-Injection for Accurate Lifetime Prediction," ECS Transactions, vol. 6, no. 3, pp. 229–243, Dec. 2019, doi: 10.1149/1.2728799.

СОЗДАНИЕ СХЕМОТЕХНИЧЕСКИХ МОДЕЛЕЙ ЭЛЕМЕНТНОЙ БАЗЫ ОТЕЧЕСТВЕННЫХ ОУ

Камаров Л.Ш., Шумкова К.Ю.

Национальный исследовательский университет
«Высшая школа экономики»
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе были рассмотрены подходы к созданию SPICE-моделей операционных усилителей (ОУ). Представлена схема измерений, позволяющая исключить эффекты самовозбуждения тестовых структур, а также приведены результаты измерения электрических характеристик тестовых прп и рпр структур, применяющихся при проектировании отечественных ОУ. По результатам измерений разработаны SPICE-модели биполярных транзисторов, позволяющие проводить схемотехническое моделирование ОУ.

Введение

Операционный усилитель является универсальным функциональным элементом, широко используемым в современных схемах формирования и преобразования информационных сигналов различного назначения как в аналоговой, так и в цифровой технике [1]. В настоящее время для проектирования различной аппаратуры широко применяются схемотехнические САПР. Однако, для более эффективного применения САПР требуется разрабатывать схемотехнические модели отечественных электрорадиоизделий (ЭРИ), т.е. их электронные двойники. Данная задача может быть решена несколькими способами. Обычно выделяют два способа: 1) макромодельный подход [2], который заключается в добавлении в схему дополнительных источников тока и напряжения для возможности моделирования электрических характеристик реального ЭРИ; 2) реализация электрической схемы ЭРИ [3] с применением SPICE-моделей полупроводниковых приборов, входящих в состав ЭРИ.

Первый подход к созданию электронных двойников ЭРИ является более простым и универсальным, так как для ее создания требуется только знание основных электрических параметров ЭРИ и не требуется его электрическая схема, не нужно создавать тестовые структуры и проводить измерения их параметров и электрических характеристик. Однако второй подход позволяет наиболее точно описать электрические характеристики ЭРИ и проводить анализ его схемотехнических решений и при необходимости выполнять его доработку для улучшения его параметров и характеристик.

Поэтому данная работа была посвящена разработке SPICE-моделей интегральных биполярных транзисторов, которые будут использоваться для проектирования отечественных ОУ и схем на их основе.

Экспериментальное исследование тестовых структур

Для разработки SPICE-моделей npn и pnp биполярных транзисторов проводились измерения статических и динамических характеристик пяти транзисторов каждого типа. Измерения статических характеристик тестовых структур биполярных транзисторов проводились с помощью прецизионного мультиметра Keithley 2602 по четырехпроводной схеме, а для измерения динамических характеристик использовались генератор АКИП-3417 и осциллограф АКИП-4134.

В процессе измерения статических электрических характеристик тестовых структур возникал паразитный эффект [4, 5], который проявлялся на входной характеристике в виде резкого увеличения тока коллектора и тока базы при напряжении база-эмиттер от 0,6 В до 0,8 В, что также негативно сказывается на выходной характеристике (см. рис. 1).

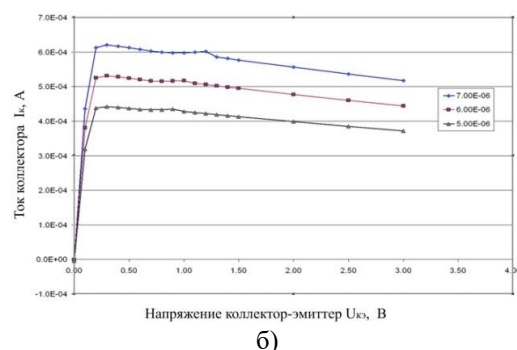
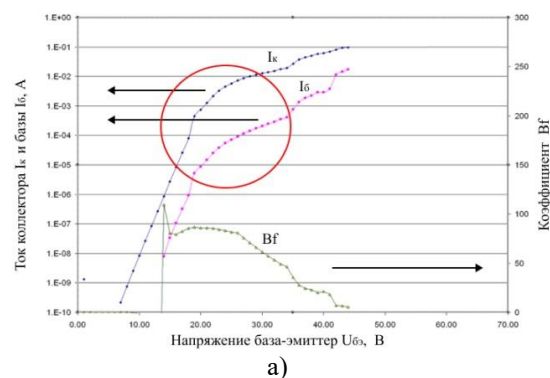


Рис. 1. Влияние паразитных эффектов на входные (а) и выходные (б) характеристики биполярного транзистора

Для устранения этого паразитного эффекта в схему измерения статических параметров были добавлены дроссели с индукцией 10 мкГн.

На рис. 2 приведена схема измерений электрических характеристик, позволяющая подавлять эффекты самовозбуждения.

На рис. 3 и 4 приведены результаты измерения статических характеристик npn транзистора, на которых отсутствуют эффекты самовозбуждения.

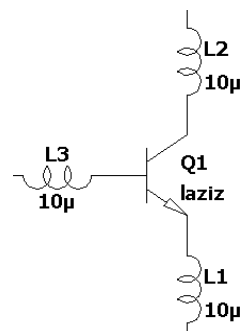


Рис. 2. Схема измерений для подавления паразитных эффектов

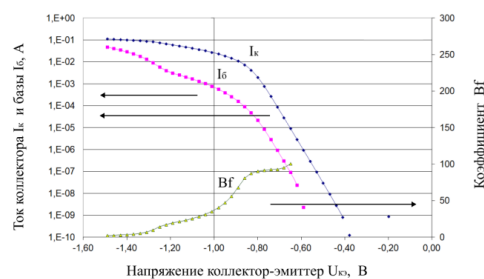


Рис. 3. Характеристика Гуммеля npn транзистора

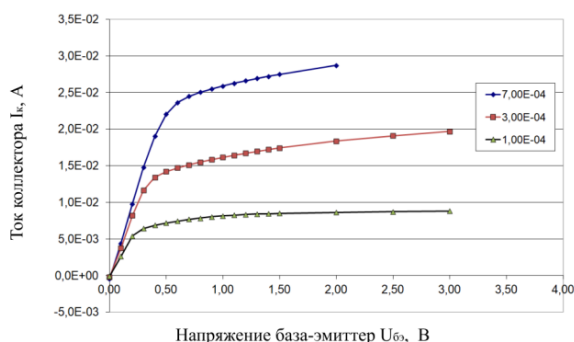


Рис. 4. Выходные характеристики при транзистора

Разработка SPICE-моделей тестовых структур

Для определения параметров SPICE-модели npn и pnp транзисторов были выбраны наиболее типичные электрические характеристики для данных тестовых структур.

По результатам измерений были определены параметры SPICE-модели Гуммеля-Пуна npn и pnp транзисторов, которые приведены в таблице 1.

Таблица 1. Основные параметры SPICE- модели Гуммеля-Пуна для npn транзистора

Параметр SPICE-модели	Значение
IS	1.038000E-16
BF	103.800000
NF	1.000000
IKF	0.018030
ISE	5.559000E-16
NE	1.570000
BR	8.872000
NR	0.995300
IKR	0.003837
ISC	2.462000E-16
NC	1.479000
RB	1.230000
RE	3.890000
RC	11.380000

Схема моделирования тестовых структур приведена на рис. 5.

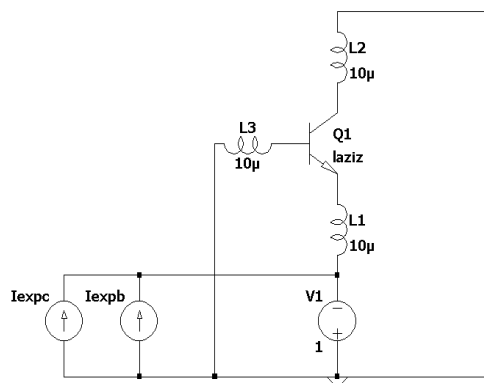


Рис. 5. Схема моделирования входной характеристики биполярного транзистора

На рис. 6 и 7 представлено сравнение результатов измерения и моделирования для npn транзистора.

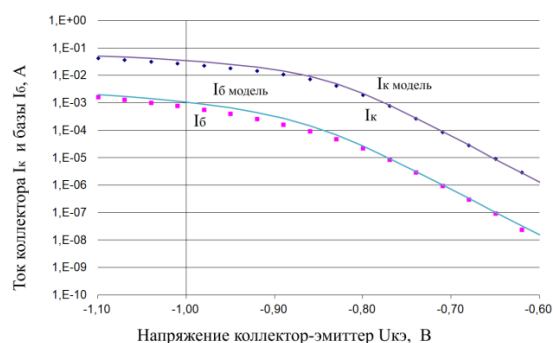


Рис. 6. Сравнение измеренных (точки) и смоделированных (линии) входных характеристик тестовых структур

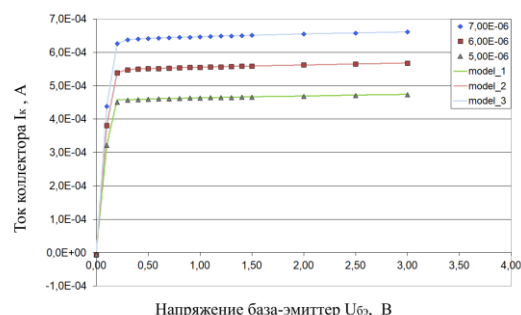


Рис. 7. Сравнение измеренных (точки) и смоделированных (линии) выходных характеристик тестовых структур

Из результатов сравнения измеренных и смоделированных характеристик тестовых структур видно хорошее совпадение. При этом расхождение смоделированных и измеренных характеристик составляет не более 15%.

Разработанные модели будут использоваться для проектирования отечественных операционных усилителей.

Заключение

В результате выполнения работы была разработана схема для измерения электрических характеристик npn и pnp биполярных транзисторов, позволяющая исключать паразитные эффекты на входной и выходной характеристиках.

По результатам измерения выполнена экстракция SPICE-параметров npn и pnp биполярных транзисторов, а также проведено сравнение экспериментальных и смоделированных данных. При этом расхождение смоделированных и измеренных характеристик составляет не более 15%.

Разработанные SPICE-модели биполярных транзисторов будут использоваться при проектировании отечественных операционных усилителей и схем на их основе.

Список литературы

1. Павлейно М. А., Ромаданов В. М., Сафронова Ю. Ф., Статия А.А. Моделирование работы электрических схем в simulink. Применение операционных усилителей для фильтрации, усиления и генерации

сигналов. Учебно-методическое пособие / Санкт-Петербург, 2007. – с. 136.

2. Азимов Э. И., Андреев В. В. Макромоделирование интегральных микросхем операционных усилителей. Электронный журнал: наука, техника и образование. 2021. № 2 (33). – с. 50-61.

3. Бакеренков А. С., Соломатин А. В., Першенков В. С. Использование схмотехнического моделирования при прогнозировании радиационной стойкости биполярных операционных усилителей // Научная сессия НИЯУ МИФИ-2014. – 2014. – с. 61-61.

4. Dvornikov O. V. et al. The main characteristics of SiGe HBTs at low temperatures // Вісник Національного технічного університету України Київський політехнічний інститут. Серія: Радіотехніка. Радіоапаратобудування. – 2016. – №. 66. – с. 87-96.

5. Ивлиев И. А., Калиновский Д. В. Учет влияния паразитных эффектов и типовых конструктивных технологических решений, используемых при проектировании РЭС. 2018. – с. 294-295.

МОДЕЛИРОВАНИЕ ЭЛЕКТРО-ТЕПЛОВОГО РЕЖИМА БЛОКА УПРАВЛЕНИЯ ЭЛЕКТРИЧЕСКИМ ДВИГАТЕЛЕМ ТОПЛИВНОГО НАСОСА

Клопотов Г.И., Ивлев К.И.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе промоделирована электрическая схема силового блока управления электрическим двигателем топливного насоса, рассчитаны тепловые параметры компонентов и мощности, выделяемые на них, составлены тепловые модели мощных элементов схемы, рассчитаны температуры компонентов.

Введение

При проектировании силовых электронных схем одной из важнейших проблем является обеспечение теплового режима компонентов схемы. Электро-тепловое моделирование схемы позволяет подробно исследовать влияние режимов работы схемы и условий охлаждения компонентов [1], [2].

Актуальность темы исследования определяется необходимостью обеспечения работоспособности электронной схемы управления путем учета и исключения перегрева элементов при проектировании. Электро-тепловой расчет и верификация блока управления необходима, чтобы убедиться в достоверности и правильности его работы для точной замены импортных компонентов на отечественные аналоги и предсказания нового теплового режима схемы. Новизна работы заключается в автоматизированном подходе к разработке и верификации электро-тепловых моделей силовых компонентов схемы управления, позволяющих значительно сокра-

тить время при составлении электро-тепловых схем и расчете теплового режима компонентов.

В настоящей работе была поставлена цель провести исследование электрических и тепловых режимов блока управления электрическим двигателем с помощью специально разработанного нами программного обеспечения, получающего из SPICE значения выделяемой мощности компонентов и составляющего тепловые модели компоненттрв. Для расчета теплового режима из нескольких возможных методов был выбран метод электротепловой аналогии как наиболее совместимый с программным обеспечением SPICE и программным обеспечением для автоматизированного создания тепловых моделей электрических компонентов.

Оценка тепловых параметров корпусов компонентов

В схеме имеются мощные компоненты: транзисторы FCA20N60F и IRFB18N50K и диоды DSEE29-12CC, размещенные в корпусах TO-3PN, TO-220AB и ISOPLUS-220 соответственно. Отвод тепла производится на алюминиевую стенку блока управления. Ниже на рис.1 приведен фрагмент схемы с выводами к электрическому двигателю топливного насоса.

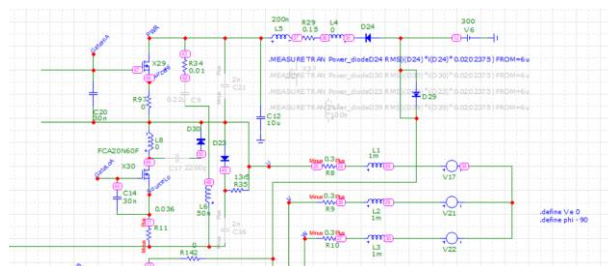


Рис. 1. Фрагмент силовой схемы блока управления электрическим двигателем

Тепловые параметры корпусов были рассчитаны стандартным способом и сравнены с тепловыми параметрами, приведенными в описания компонентов.

Исследование электрического и теплового режимов работы схемы управления с помощью программы схмотехнического моделирования Micro-sap

Использовалось написанное авторами программное обеспечение для автоматизированного получения из SPICE выделяемой на электрических элементах мощности и создания нетлиста тепловой модели по имеющимся в разработанной базе данных тепловым моделям. Далее, с помощью программы схмотехнического моделирования Micro-sap анализировался тепловой режим схемы с учетом тепловых параметров корпусов компонентов. В результате проводился анализ перегрева элементов. Полученные мощности компонентов передавались далее в нашу программу для составления тепловых моделей каждого из выбранных элементов схемы. В результате были получены графики зависимости температуры компонентов от времени.

На рис. 2 приведены расчетные мощности транзистора IRFB18N50K и диода DSEE29-12CC.

Function	✓	1
Power_DSEE29_12CC		62.676
Power_IRFB18N50K		41.185

Рис. 2. Мощности элементов DSEE29-12CC и IRFB18N50K

Полученные мощности были добавлены в собранную тепловую модель с учетом наиболее неблагоприятных условий: температура наружной стенки корпуса составляет 80 °С (рис. 3). Так как тепловой режим для остальных компонентов совпадает, то на рисунке приведена только часть из них, полностью отражающая тепловые характеристики для мощных компонентов.

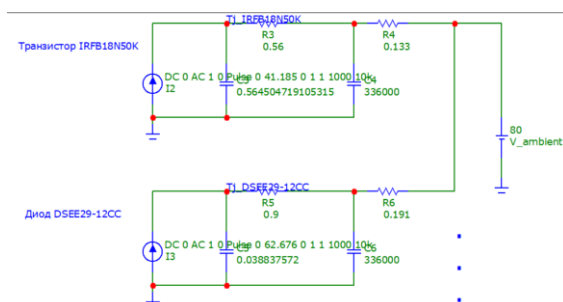


Рис. 3. Тепловая модель

В результате с помощью тепловой схемы были получены графики зависимости температуры компонентов от времени ($t=1000$ секунд) для диода и транзистора (рис. 4 и рис. 5)

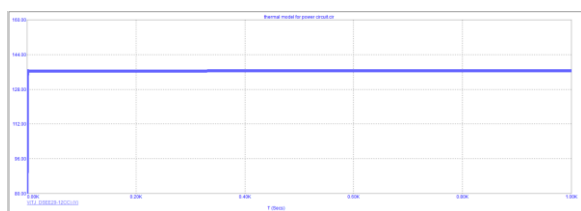


Рис. 4. График зависимости температуры диода DSEE29-12CCB от времени

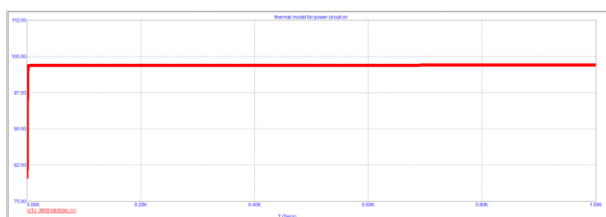


Рис. 5. График зависимости температуры транзистора IRFB18N50K от времени

Заключение

Подводя итоги, стоит отметить, что в данной работе были автоматизировано получены тепловые модели электрических компонентов, составляющие основу для формирования единой базы данных для распространенных типов корпусов силовых элементов.

С использованием тепловых моделей для мощных электрических компонентов схемы управления был проведен анализ теплового режима работы схемы. Для диода DSEE29-12CC после установления теплового режима температура составила около 136 °С, для транзистора IRFB18N50K - 103 °С

Кроме того, можно сделать вывод о практической применимости полученных программы и результатов моделирования при разработке других электрических схем и приборов, в частности, при расчете тепловых нагрузок отдельных элементов схемы.

В заключении хотелось бы выделить основные направления для дальнейшего развития. Во – первых, актуально наполнение базы данных различными типами компонентов помимо силовых. Во – вторых, расчет тепловых моделей отечественных корпусов в силу отсутствия информации о тепловых параметрах в большей части документации производителей.

Список литературы

1. Г. Н. Дульнев, Н. Н. Тарновский. Тепловые режимы электронной аппаратуры. «Энергия», 1971. – 249 с.
2. А. И. Колпаков. Особенности теплового расчета импульсных силовых каскадов. «Компоненты и технологии, № 1.2002. – 5 с.
3. Володин В. Я. LTspice: компьютерное моделирование электронных схем, 2020. – 392 с.
4. БОЛДЫРЕВ А. В. МОДЕЛИРОВАНИЕ ЭЛЕКТРОННЫХ СХЕМ В ПРОГРАММЕ MICRO-CAP. – 2021.

ВАРИАНТЫ МОДЕРНИЗАЦИИ МОДУЛЯ СПЕЦВЫЧИСЛИТЕЛЯ БЕСПЛАТФОРМЕННОЙ ИНЕРЦИАЛЬНОЙ НАВИГАЦИОННОЙ СИСТЕМЫ

Беликов И.Д.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе предложены варианты модернизации бесплатформенной инерциальной навигационной системы (БИНС): замена элементной базы модуля спецвычислителя (МСВ) и изменение уровня сигналов технологических интерфейсов, доработка алгоритмов обработки информации с датчиков в соответствии с требованиями по увеличению быстродействия или точности выходных данных системы. Подобная модернизация позволит увеличить вычисли-

тельную способность МСВ, повысив точность выходной информации, и добавить вариативность применения БИНС в различных системах

Введение

Инерциальные системы измеряют угловые скорости и ускорения объекта относительно его трех осей. При обработке этих данных появляется возможность определить местоположение объекта относительно его начальных координат. Ключевой особенностью БИНС является отсутствие гиросtabilизированной платформы, которое значительно уменьшает массу и габариты системы, но добавляет проблемы, связанные с накапливающейся ошибкой. БИНС позволяет объекту попасть в точку назначения, зная лишь начальные и конечные координаты, без использования внешних систем связи [1]. Блок электроники рассматриваемой БИНС содержит в своем составе модуль спецвычислителя, который отвечает за обработку поступающей с датчиков информации и вычисление корректирующих сигналов для бортовой системы управления (БСУ) объекта. Существует необходимость постоянного совершенствования средств, обеспечивающих ориентацию в пространстве высокоманевренных объектов. Из-за устаревания и отсутствия возможности закупки некоторых электронных компонентов требуется замена их аналогами. Также необходима модернизация алгоритмов обработки информации для уменьшения накапливающейся ошибки навигации.

В данной работе рассматриваются несколько способов повышения производительности МСВ БИНС.

Особенности работы МСВ

Модуль спецвычислителя выполняет ряд важных функций, обеспечивающих работу БИНС. Среди них такие как: общение с периферийными модулями по технологическим интерфейсам (получение синхросигналов с модуля формирования синхроимпульсов (МФСИ), получение акселерометрических данных с модуля аналого-цифрового преобразователя (МАЦП), получение данных об угловой скорости с волоконно-оптических гироскопов (ВОГ), опрос приемника глобальной навигационной спутниковой системы (ГНСС), опрос оптических систем коррекции), обработка и комплексирование поступающей информации с внешних датчиков и систем и вывод навигационных данных в БСУ объекта. В состав МСВ входят 2 микропроцессора Миландр 1986ВЕ1Т, которые контролируют работу системы, и интерфейсные шины взаимодействия с другими модулями. Например, коммуникация МСВ с МАЦП происходит через шину ISA, для которой требуется не только протокол сбора и отправки данных, но и управление конверторами физических уровней шины. Прием данных с триады ВОГ осуществляется по интерфейсу RS-485 с тактированием через внешний синхроимпульс. Через модуль мультиплексного канала информационного обмена (МКИО) происходит общение БИНС с другими системами, его работа обеспечивается модулем микропроцессора по ГОСТ 52070-2003[2], который работает в широком диапазоне

температур от -60°C до $+60^{\circ}\text{C}$ и сертифицирован для применения в авиационной технике.

Модификации аппаратной части

Микропроцессоры, которые в данный момент стоят в МСВ, перестали выпускаться в связи с введенными по отношению к РФ санкциями, поэтому требуется найти замену у отечественного производителя.

В МСВ старого образца используются однокядерные микропроцессоры на базе вычислительного ядра RISC V. Среди отечественных производителей есть предложения с микропроцессорами, работающими на тех же частотах, но с несколькими ядрами ARM M4. Использование многоядерных микропроцессоров позволит разделить выполнение алгоритмов обработки и ускорить вычисление. Предложенная модификация позволит увеличить быстродействие модуля в несколько раз.

В связи с тем, что все составные модули блока электроники БИНС проектируются индивидуально для конкретного прибора, есть предложение отказаться от стандартизации шины ISA и исключить преобразователи уровней напряжения из 3.3В в 5В и наоборот. Такое изменение позволит увеличить энергоэффективность и повысить надежность системы уменьшением числа электронных компонентов, а также увеличить скорость быстродействия шины, исключив задержку преобразования и упростив исполняемый код.

Модификации программной части: параллельные расчеты при большей частоте синхросигналов

Алгоритмы сбора информации, поступающей с датчиков, в основном останутся неизменными, но в связи с заменой микропроцессоров на более современные, можно увеличить быстродействие системы, разбив их выполнение на разные ядра. Реализовать параллельные вычисления можно реализовать в соответствии MPI и OpenMP, на архитектуре новых процессоров[3]. В данный момент обработка информации ведется по внешним тактовым сигналам, в течение периода которых необходимо провести все вычисления и отправить данные на другой процессор или в БСУ объекта. При разбиении алгоритмов на параллельные процессы, выполняемые на разных ядрах, можно увеличить частоту сигнала, по которому начинается обработка поступающей информации. С увеличением частоты синхросигналов вырастет количество получаемой информации за промежуток времени, что позволит быстрее реагировать на возмущающие воздействия.

Модификации программной части: параллельные расчеты при усложнении алгоритмов обработки

С другой стороны, можно усложнить алгоритмы обработки и комплексирования навигационных данных, тем самым повысив их точность. При сохранении тактов синхроимпульсов будет возможна реали-

зация повышения точности алгоритмов за счет специальной организации вычислений, при которой процесс определения ориентации идет параллельно на нескольких ядрах. Повышая вычислительную мощность микропроцессоров, появится возможность внедрить алгоритм оптической астро-инерциальной навигации. Он работает по принципу получения изображения звездного неба, вычисления взаимных координат выявленных звезд и его сравнения с эталонным каталогизированным значением, заложенным в памяти прибора[4]. Помимо звездного датчика появится возможность включения в работу модуля ГНСС[5], который позволит в некоторые промежутки времени получить координаты объекта со спутника и проводить анализ ошибок инерциальной системы.

Реализация алгоритмов навигации БИНС на нескольких ядрах микропроцессора в совокупности с алгоритмами навигации по звездному датчику и ГНСС позволит в реальном времени производить коррекцию навигационных данных, путем сравнения и/или комплексирования данных с независимых источников.

Заключение

Приведенные в данной работе варианты модернизации МСВ БИНС позволят увеличить скорость реакции на возмущающее воздействие внешних факторов. Замена элементной базы позволит увеличить вычислительную мощность системы, повысить энергоэффективность и увеличить надежность. Внесение корректировок в алгоритмы обработки информации и интерфейсного общения модулей позволит увеличить общую точность навигации и/или скорость реакции системы. Итогом модернизации станет повышение конкурентоспособности отечественных систем инерциальной навигации над зарубежными аналогами.

Список литературы

1. Матвеев В.В., Распопов В.Я. Основы построения бесплатформенных инерциальных навигационных систем. - СПб: ГНЦ РФ ОАО «Концерн «ЦНИИ «Электронприбор», 2009. - 280 с.
2. ГОСТ Р 52070-2003. Интерфейс магистральный последовательный системы электронных модулей. Общие требования : национальный стандарт Российской Федерации : дата введения 2004-01-01 / Главное управление технической политики в области стандартизации Госстандарта России. – Изд. Официальное. – Москва: ИПК Издательство стандартов, 2003. -22 с.
3. Шпаковский Г.И. Реализация параллельных вычислений: кластеры, грид, многоядерные процессоры, квантовые компьютеры. - Минск: БГУ, 2010. - 155 с.
4. Бессонов Р.В., Белинская Е.В., Брысин Н.Н., Воронков С.В., Куркина А.Н., Форш А.А. Звёздные датчики ориентации в астроинерциальных системах летательных аппаратов // Современные проблемы дистанционного зондирования Земли из космоса. - 2018 Т. 15. - № 6. - С. 9-20.
5. Соловьев Ю.А. Системы спутниковой навигации. - М.: Эко-Трендз, 2000. - 270 с.

МОДЕЛИРОВАНИЕ InGaAs/InAlAs ОДНОФОТОННЫХ ЛАВИННЫХ ФОТОДИОДОВ С ЦЕЛЮ ОЦЕНКИ ВЛИЯНИЯ ДИАМЕТРА АКТИВНОЙ ОБЛАСТИ НА УРОВЕНЬ ТЕМНОВОГО СЧЕТА

Филяев А.А.^{1,2,3}, Заводиленко В.В.^{1,2,3},
Павлов И.Д.^{2,3}, Ожегов Р.В.¹

¹Национальный исследовательский университет
«Высшая школа экономики»,
Базовая кафедра квантовой оптики и
телекоммуникаций ЗАО "Сконтел",
МИЭМ НИУ ВШЭ;

²Национальный исследовательский
технологический университет МИСиС,
Лаборатория квантовых коммуникаций;

³ООО «КуРэйт»

Аннотация

В работе рассматривается влияние диаметра активной области на уровень темнового счета (DCR) InGaAs/InAlAs однофотонных лавинных фотодиодов (ОЛФД). Данное исследование проводится в рамках оптимизации конструкционных параметров ОЛФД. Моделирование диодной структуры проводилось в расчетной среде T-CAD. Установлено, что при уменьшении диаметра активной области, снижается уровень темного тока, а это в свою очередь приводит к уменьшению DCR при работе устройства в Гейгеровском режиме.

Введение

Однофотонный лавинный фотодиод (ОЛФД) для детектирования излучения с длиной волны $\lambda = 1550$ нм может быть выполнен из различных материалов. В научном сообществе продолжается развитие ОЛФД на основе таких материалов, как Si/Ge, InGaAs/InP и InGaAs/InAlAs [1–12]. В настоящей работе будет проведено исследование такой перспективной структуры, как InGaAs/InAlAs.

Преимущество материала InAlAs в качестве зоны умножения заключается в большой подвижности электронов, что позволяет значительно быстрее гасить возбужденную лавину и переводить структуру в равновесное состояние. Таким образом, данное устройство может обладать большей предельной частотой работы и лучшими характеристиками по послепульсам, в сравнении с устройством с зоной умножения, выполненной из материала InP. Это обусловлено тем, что подвижность дырок в InP ниже, чем подвижность электронов в InAlAs [13].

Тем не менее, устройства на рынке ОЛФД для $\lambda = 1550$ нм представлены в основном на материалах InGaAs/InP. Основным препятствием перед созданием ОЛФД на основе материалов InGaAs/InAlAs является большое количество дефектов в материале InAlAs, что не позволяет использовать весь его потенциал [14, 15]. Улучшение технологических процессов напыления данного материала делает возможным конкурентное двух данных типов устройств. Оптимизация структуры является второй

важной задачей раскрытия потенциала устройства InGaAs/InAlAs.

В настоящей работе проведено исследование влияния диаметра активной области на уровень темнового счета InGaAs/InAlAs ОЛФД с целью оптимизации его конструктивных параметров.

Построение 2D-модели

Моделирование производилось в системе T-CAD. При моделировании в T-CAD конкретные значения времени жизни носителя заряда задавались в качестве параметра для расчета скорости рекомбинации по механизму Шокли-Рида-Холла. Также в расчете учитывались излучательная и Оже рекомбинации. При моделировании методом конечных элементов решалась система уравнений, состоящая из уравнений непрерывности для электронов и дырок, и уравнения Пуассона.

Для того, чтобы ограничить активную область зоны умножения (область, где происходит основной процесс лавинной генерации и поглощения фотона), необходимо использовать разные ширины области умножения в активной и не активной области умножения. В активной области необходимо достичь большей напряженности поля, следовательно, ширина запрещенной зоны должна быть меньше.

В рамках данного исследования рассматривалась структура с тремя уровнями зоны умножения с плавным переходом (см. рис. 1).

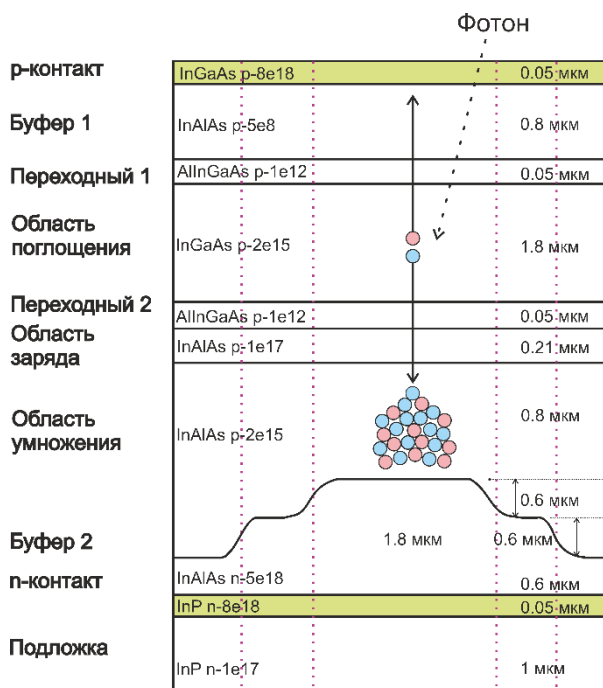


Рис. 1. Схематическое представление 2D-структуры ОЛФД с тремя уровнями зоны умножения и плавным переходом

Основным преимуществом такой структуры является низкое значение локального повышения напряженности электрического поля (см. рис. 2).

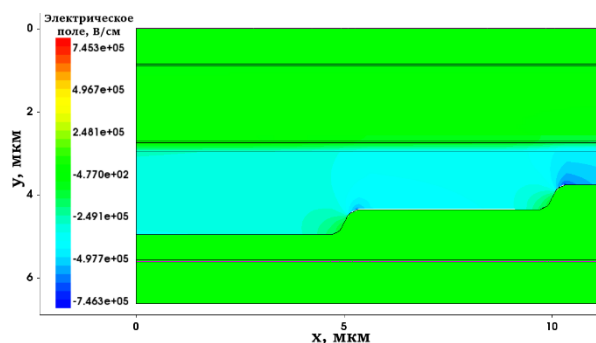


Рис. 2. Тепловая карта распределения напряженности электрического поля в структуре с тремя уровнями зоны умножения с плавным переходом

Такой эффект приводит к обеспечению большего значения квантовой эффективности ОЛФД при его эксплуатации.

Моделирование устройств с различным диаметром активной области

Моделирование проводится для структур с диаметром 25 мкм, 15 мкм и 10 мкм. Основным преимуществом использования структур с уменьшенным диаметром активной области является меньшее значение темнового тока. Однако, основным недостатком является необходимость дополнительной фокусировки попадающего на диод излучения. В представленных результатах моделирования было проведено сравнение вольтамперных характеристик (ВАХ) темнового тока для диодов с различным диаметром активной области (см. рис. 3).

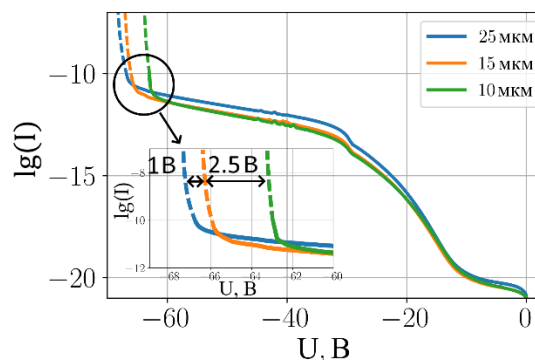


Рис. 3. Сравнение темнового ВАХ для ОЛФД со структурой с тремя уровнями зоны умножения и плавными переходами для различных диаметров активной области

Как видно на данном графике, на линейном участке ВАХ меньшим темновым током обладают устройства с меньшим диаметром активной области. Однако у данных структур аналогично меньше напряжение пробоя.

Темновой ток на линейном участке для структуры с диаметром 25 мкм выше, чем для структуры с диаметром 10 мкм приблизительно в 3 раза. Темновой ток на линейном участке приблизительно равен для структуры с диаметром 15 мкм и 10 мкм.

Напряжение пробоя для структуры с диаметром активной области 25 мкм выше, чем напряжение пробоя структуры с диаметром 15 мкм на 1 В, у которой, в свою очередь, напряжение пробоя выше, чем напряжение пробоя у структуры с диаметром 10 мкм на 2.5 В.

Заключение

В результате исследования было установлено, что изменение диаметра активной области структуры приводит как к изменению темнового тока на линейном участке ВАХ, так и к изменению напряжения пробоя. Было продемонстрировано, что уменьшение диаметра с 25 мкм до 10 мкм позволило снизить темновой ток в линейном режиме работы устройства приблизительно в 3 раза, однако привело к уменьшению напряжения пробоя приблизительно на 3,5 В. Первое свойство приводит к уменьшению DCR при работе устройства в Гейгеровском режиме. Однако второе свойство уменьшает вероятность детектирования фотона.

Список литературы

1. Signorelli F. et al. Low noise InGaAs/InP SPAD for fiber-based quantum applications //2021 IEEE Photonics Conference (IPC). – IEEE, 2021. – С. 1-2.
2. Liang Y. et al. Low-Timing-Jitter GHz-Gated InGaAs/InP Single-Photon Avalanche Photodiode for LIDAR //IEEE Journal of Selected Topics in Quantum Electronics. – 2021. – Т. 28. – №. 2. – С. 1-7.
3. Kizilkan E. et al. Guard-Ring-Free InGaAs/InP Single-Photon Avalanche Diode Based on a Novel One-Step Zn-Diffusion Technique //IEEE Journal of Selected Topics in Quantum Electronics. – 2022. – Т. 28. – №. 5. – С. 1-9.
4. Robinson M. B. et al. Diffuse correlation spectroscopy beyond the water peak enabled by cross-correlation of the signals from InGaAs/InP single photon detectors //IEEE Transactions on Biomedical Engineering. – 2021. – Т. 69. – №. 6. – С. 1943-1953.
5. Baek S. H. et al. Room temperature quantum key distribution characteristics of low-noise InGaAs/InP single-photon avalanche diode //Journal of the Korean Physical Society. – 2021. – Т. 78. – №. 7. – С. 634-641.
6. Kirdoda J. et al. Geiger mode Ge-on-Si single-photon avalanche diode detectors //2019 IEEE 2nd British and Irish Conference on Optics and Photonics (BICOP). – IEEE, 2019. – С. 1-4.
7. Wanitzek M. et al. GeSn-on-Si Avalanche Photodiodes for Short-Wave Infrared Detection //ESSCIRC 2022-IEEE 48th European Solid State Circuits Conference (ESSCIRC). – IEEE, 2022. – С. 169-172.
8. Lee Y. S. et al. Neat temporal performance of InGaAs/InAlAs single photon avalanche diode with step-wise electric field in multiplication layers //IEEE Access. – 2021. – Т. 9. – С. 32979-32985.
9. Chang S. C., Lee Y. S. Investigation of Breakdown Voltage Characteristics of InGaAs/InAlAs Single Photon Avalanche Diodes //2019 24th Optoelectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC). – IEEE, 2019. – С. 1-3.
10. Wu P. L. et al. Improved Performance of InGaAs/InAlAs Single Photon Avalanche Diode with Dual Multiplication Layers //Optical Sensors. – Optical Society of America, 2021. – С. STu1B. 4.
11. Zhang J. et al. Triple-Mesa InGaAs/InAlAs Single-Photon Avalanche Diode Array for 1550 nm Photon Detection //CLEO: Science and Innovations. – Optica Publishing Group, 2021. – С. SM1D. 6.
12. Zhang J. et al. Heterogeneously integrated InGaAs/InAlAs SPAD on SOI photonics platform //Advanced Photon Counting Techniques XVI. – SPIE, 2022. – С. PC120890F.
13. Chen J. et al. Study on impact ionization in charge layer and multiplication layer of InAlAs/InGaAs SAGCM avalanche photodiodes //Optoelectronic Devices and Integration VII. – SPIE, 2018. – Т. 10814. – С. 289-295.
14. Jiang Y., Chen J. Optimization of the Linearity of InGaAs/InAlAs SAGCM APDs //Journal of Lightwave Technology. – 2019. – Т. 37. – №. 14. – С. 3459-3464.
15. Sim J. S. et al. Optimization of charge and multiplication layers of 20-Gbps InGaAs/InAlAs avalanche photodiode //ETRI Journal. – 2021. – Т. 43. – №. 5. – С. 916-922.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ФРАКТАЛЬНОГО МЕТАМАТЕРИАЛА НА ЭЛЕКТРОДИНАМИЧЕСКИЕ ПАРАМЕТРЫ RFID-АНТЕННЫ

Башкевич С.В., Лаврухин И.Р.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Данная работа направлена на исследование фрактального метаматериала и его влияния на электродинамические характеристики антенны, использующейся в системах радиочастотной идентификации (RFID). В рамках данной работы предложена и исследована фрактальная структура, основанная на квадрате Серпинского. Для исследования влияния предложенной структуры на характеристики антенны была взята существующая RFID-антенна с известными параметрами. Полученные в ходе исследования характеристики коэффициента отражения S11 и диаграммы направленности антенны с метаматериалом позволяют сделать вывод о возможности применения предложенной структуры в системах RFID.

Введение

Метаматериалы – это искусственно сформированные и особым образом структурированные среды, обладающие электромагнитными свойствами, сложно достижимыми технологически либо не встречающимися в природе. Все искусственно созданные или естественные среды можно разделить на группы в зависимости от показателей их диэлектрической и

магнитной проницаемостей – ϵ и μ соответственно. Почти все встречающиеся материалы на Земле имеют положительное значение данных показателей. Материалы, имеющие отрицательное значение одного (singlenegative - SNG) или сразу обоих показателей (dualnegative - DNG), обладают нетипичными свойствами. Так, например, DNGматериалы обладают отрицательным коэффициентом преломления[1]. Метаматериалы находят широкое применение в антенной технике. Применение метаматериалов в виде подложек или излучателей антенн позволяет добиться: миниатюризации устройств, увеличение полосы пропускания, а также изменение диаграммы направленности.

Термин фрактал ввел математик и основатель фрактальной геометрии Бенуа Мандельброт в 1975 году. Образованное от латинского причастия fractus, данное слово означает: разбитый, разломанный[2]. Фрактал по своей сути является множеством обладающим свойством самоподобия, то есть целое имеет форму одной или нескольких своих частей. Геометрические формы фракталов также находят применение в дизайне антенн. В работе [3] приводятся различные примеры фрактальных антенн, а также отмечаются преимущества использования фрактальных структур. Например, самоподобие бывает полезно для антенн, работающих в нескольких частотных диапазонах, а самозаполнение позволяет снижать реальные размеры антенны при сохранении ее электрической длины. Также отмечается, что фрактальные поверхности можно использовать при построении антенн высокой направленности.

И фракталы, и метаматериалы находят свое применение в антенной технике благодаря своим нетипичным свойствам. В этой работе предлагается исследование, находящееся на стыке применения обоих направлений. Целью исследования ставится проверка гипотезы о возможности использования метаматериала на основе фрактала в системах RFID.

Исследуемая поверхность

Фракталы могут использоваться для дизайна метаматериалов или высокоимпедансных поверхностей. В данной работе рассматривается фрактальная структура, называемая ковер (квадрат) Серпинского. Несколько итераций построения квадрата Серпинского представлены на рисунке 1.

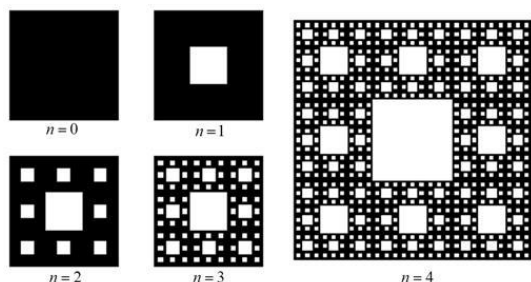


Рис. 1. Построение структуры

На первом шаге строится квадрат со стороной равной $\frac{1}{3}$ стороны первоначального квадрата, на вто-

ром -восемь квадратов со стороной равной $\frac{1}{9}$ и так далее. В оригинале площадь полученных квадратов вырезается из общего полотна. В данной работе каждый сформированный описанным выше методом квадрат будет являться излучателем, наложенным на подложку антенны. Исследуемая поверхность, смоделированная в программе CSTStudioSuite, представлена на рисунке 2.

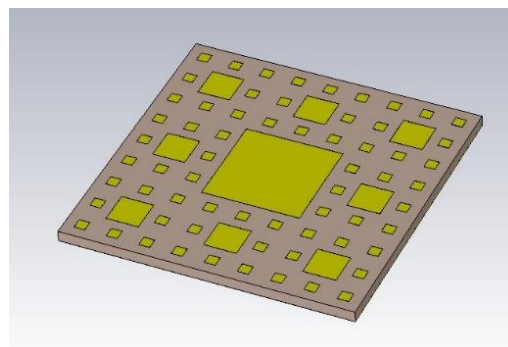


Рис. 2. Исследуемая структура

Для построения структуры произведено три итерации. В качестве материала для излучателей и экрана под подложкой выступает алюминий, в качестве диэлектрика подложки оксид алюминия. Для определения конечных размеров структуры рассматривается фазовая характеристика коэффициента отражения по аналогии с работами [4], [5]. На рисунках 3 и 4 приведено влияние изменения размеров излучателей и толщины подложки соответственно на фазовую характеристику коэффициента отражения.

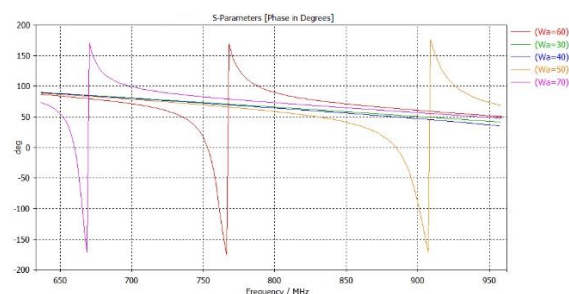


Рис. 3. Влияние размера излучателей

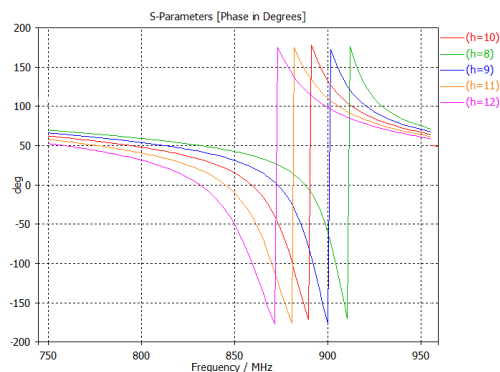


Рис. 4. Влияние толщины подложки

Рабочая полоса частот определяется в диапазоне изменения фазы от -90 до +90 градусов. Для систем RFID UHF диапазона рабочей считается полоса от 860

до 960 МГц. Скорректированные значения размера излучателя и толщины подложки равны 50 и 8 миллиметрам соответственно. Итоговый размер структуры, приведенной на рисунке 2 равен $150 \times 150 \times 8 \text{ мм}^3$

Влияние исследуемой поверхности на RFID-антенну

Для оценки влияния предложенной структуры на электродинамические характеристики RFID-антенны была взята существующая антенна, исследуемая в работе [6]. На рисунке 5 сверху представлена оригинальная антенна, а на рисунке 5 снизу с исследуемой структурой в виде подложки. Габаритные размеры антенны с метаматериалом $300 \times 300 \times 30 \text{ мм}^3$.

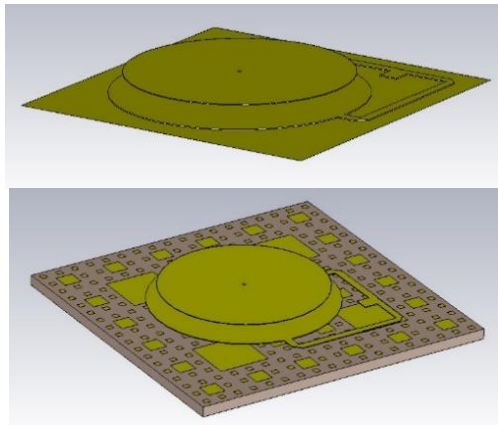


Рис. 5. Оригинальная и исследуемая антенны

Варьируя далее толщину подложки и расстояние от исследуемой структуры до антенны можно получить различного рода улучшения электродинамических характеристик антенны. Так, например, при толщине подложки в 10 миллиметров и расстоянии от нее до антенны в 2 миллиметра, заметное улучшение происходит в усилении антенны. На рисунке 6 представлена диаграмма направленности оригинальной антенны, а на рисунке 7 исследуемой.

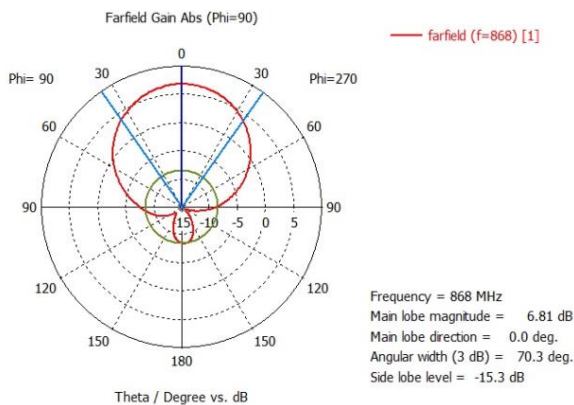


Рис. 6. ДН оригинальной антенны

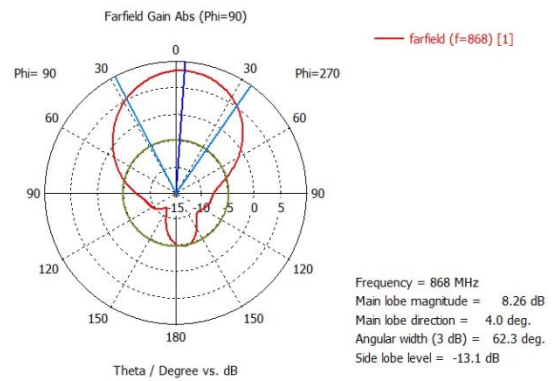


Рис. 7. ДН исследуемой антенны

Усиление антенны в направлении главного лепестка увеличилось почти на 2 децибела, при незначительном сокращении угла излучения по уровню падения мощности в 3 децибела.

При увеличении толщины подложки до 22 миллиметров, и расстояния до антенны до 5 миллиметров, происходит обратное. Незначительно падает мощность антенны в направлении главного лепестка, однако почти в 2 раза увеличивается ширина угла по уровню падения мощности в 3 децибела. Сравнение диаграмм направленности приведено на рисунке 8.

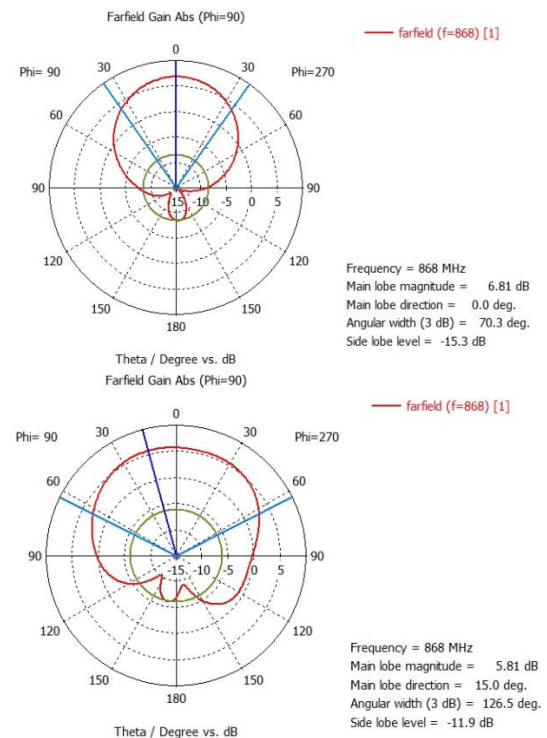


Рис. 8. ДН оригинальной (сверху) и исследуемой антенны (снизу)

Помимо изменений в диаграмме направленности антенны также изменяется и значение параметра S11. На рисунке 9 представлена характеристика коэффициента отражения для оригинальной антенны в рабочем диапазоне частот.

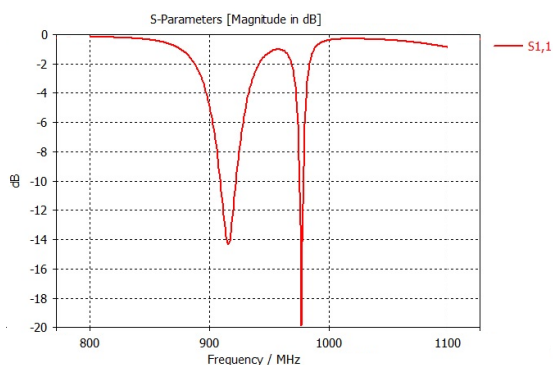


Рис. 9. Параметр S_{11} оригинальной антенны

Для антенны с толщиной подложки в 10 миллиметров и 10 миллиметровым расстоянием до антенны характеристика коэффициента отражения принимает следующий вид (Рис. 10).

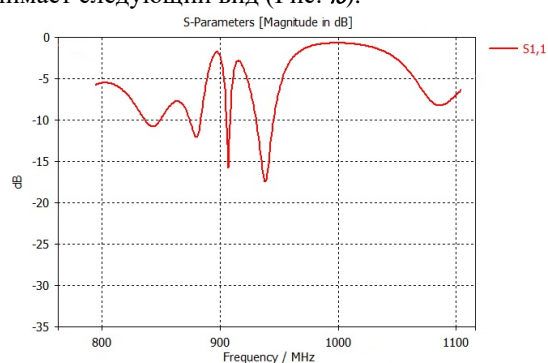


Рис. 10. Параметр S_{11} исследуемой антенны

Теперь в рабочем диапазоне частот наблюдается несколько резонансов, а удовлетворительное согласование (от -10 децибел) можно заметить на 4 полосах частот.

Для антенны с толщиной подложки в 22 миллиметра и 5 миллиметровым расстоянием характеристика коэффициента отражения принимает следующий вид (Рис. 11).

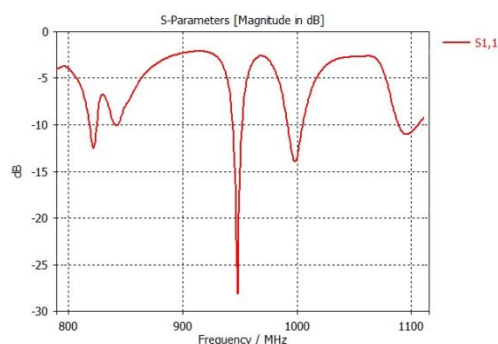


Рис. 11. Параметр S_{11} исследуемой антенны

В данном случае также наблюдается улучшение согласования в рабочем диапазоне частот.

Заключение

В результате работы была исследована структура метаматериала основанная на фрактале. По итогу работы можно сделать вывод о возможности исполь-

зования такой структуры для антенн RFID-систем. При различных комбинациях толщины подложки и расстояния от подложки до антенны наблюдаются различные улучшения электродинамических параметров антенны.

Публикация подготовлена в результате исследования (№23-00-003 «Исследование технологий и устройств беспроводного трансфера электромагнитной энергии для высокоскоростных мобильных и нательных устройств Интернета вещей (IoT/IIoT) и киберфизических систем») в рамках Программы «Научный фонд Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ)».

Список литературы

1. Slyusar, Vadym. (2010). Метаматериалы в антенной технике: основные принципы и результаты. 44 - 60.
2. The Fractal Geometry of Nature. Book by Benoit Mandelbrot.
3. W. J. Krzysztofik, "Fractals in Antennas and Metamaterials Applications," Fractal Analysis - Applications in Physics, Engineering and Technology, Jun. 2017, doi: 10.5772/intechopen.68188.
4. Hocini, Abdesselam&Melouki, Nouredine&Denidni, Tayeb. (2020). Modeling and simulation of an antenna with optimized AMC reflecting layer for gain and front-to-back ratio enhancement for 5G applications. Journal of Physics Conference Series. 1492. 12006. 10.1088/1742-6596/1492/1/012006.
5. Askari, Hussain & Hussain, Niamat& Choi, Domin&Sufian, Md. Abu & Abbas, Anees& Kim, Nam. (2021). An AMC-Based Circularly Polarized Antenna for 5G sub-6 GHz Communications. Cmc - TechSciencePress-. 69. 2997-3013. 10.32604/cmc.2021.018855.
6. I. R. Lavrukhin, S. V. Bashkevich, V. A. Tarantsev and A. A. Yelizarov, "Development of Antenna Module for the Intelligent RFID System," 2020 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Svetlogorsk, Russia, 2020, pp. 1-4, doi: 10.1109/SYNCHROINFO49631.2020.9166056.

РАЗРАБОТКА АНАЛОГОВОЙ ЧАСТИ СПЕКТРОМЕТРИЧЕСКОГО ДЕТЕКТОРА, РЕГИСТРИРУЮЩЕГО ИОНИЗИРУЮЩИЕ ИЗЛУЧЕНИЯ

Сетдигов Р.М., Тюрин И.Г.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии,
МИЭМ НИУ ВШЭ

Аннотация

В данной статье представлено исследование, моделирование и практическая реализация аналоговой части спектрометрического датчика регистрации

ионизирующего излучения на основе сцинтиллятора в сочетании с фотоэлектронным умножителем (ФЭУ). Основной проблемой является получение максимального значения каждого импульса, полученного от ФЭУ. В качестве решения был предложен подход с замедлением сигнала для диагностики достижения максимума. В результате работы была получена рабочая модель устройства, которая будет использоваться для завершения разработки портативного, автономного спектрометра ионизирующего излучения. Успешная сборка реальной схемы и ее испытание с использованием лабораторного источника сигнала подтверждают потенциал данного подхода и технологии.

Введение

Предметом данной статьи является исследование, моделирование и практическая реализации аналоговой части спектрометрического датчика регистрирующего ионизирующее излучения на основе сцинтиллятора [1] в сочетании с фотоэлектронным умножителем (ФЭУ англ. SiPM) [2] (см. рис. .). Регистрация ионизирующих излучений – важная задача, так как ионизирующее излучение оказывает негативное воздействие на живые организмы.

Сцинтилляционные детекторы являются одним из наиболее распространенных типов детекторов для регистрации ионизирующего излучения. Их применение в различных областях науки и техники, таких как ядерная физика, медицинская диагностика и экологический мониторинг, обусловлено высокой чувствительностью, быстродействием и возможностью анализа энергии частиц [3, 4]. В последние годы активно развивается направление создания компактных и автономных спектрометров на основе сцинтилляционных детекторов [5, 6].

В мировой науке существует множество работ, посвященных изучению и усовершенствованию аналоговых схем для сцинтилляционных детекторов [7, 8]. В этой статье мы стремимся представить альтернативный подход к решению задачи получения максимального значения каждого импульса, полученного от ФЭУ, исследуя возможности использования замедления сигнала для диагностики пика.

Высокочастотное ионизирующее излучение представляет собой электромагнитное поле (гамма-частицы) либо поток элементарных частиц (нейтронов, протонов, бета-, альфа-частиц).

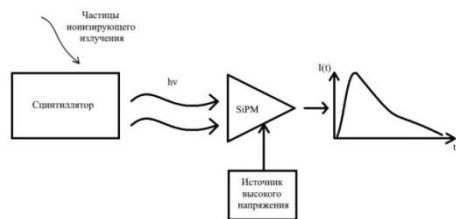


Рис. 1. Структурная схема датчика «Сцинтиллятор-ФЭУ»

Ионизирующая частица (например, электрон), проходя через сцинтиллятор, переводит атомы веще-

ства в возбужденное состояние, при возвращении в своё первоначальное состояние частицы начинают испускать фотоны света, т.е. генерируется импульс света (происходит сцинтилляция). При взаимодействии нейтральных частиц (γ -квантов) с атомами сцинтиллятора образуются вторично заряженные частицы, которые также вызывают свечение. На следующем этапе световой сигнал сцинтиллятора приходит на кремниевый фотоумножитель (далее ФЭУ) – устройство, представляющее собой матрицу из лавинных фотодиодов, каждый из которых генерирует электрический сигнал, при попадании на него фотона света. Этот сигнал с помощью созданной нами аналоговой части преобразуется к удобному виду, а далее обрабатывается.

Варианты решения задачи

Для анализа ионизирующего излучения нам необходимо каким-либо способом получить максимальное значение каждого импульса, полученного с ФЭУ, так как именно максимум пропорционален энергии ионизирующей частицы.

Получение максимума возможно следующими способами.

Частая оцифровка сигнала не подходит, так как нужны дорогостоящие быстродействующие аналого-цифровые преобразователи (АЦП), а также придется обрабатывать огромные массивы данных.

По площади под графиком импульса можно также определить максимум сигнала, но так как аналоговые шумы значительны, то схема с интегрирующим звеном не подходит, так как суммирует и шумы.

При переходе через максимальную точку производная функции сигнала будет менять знак, однако дифференцирующие цепочки дифференцируют и шумы, поэтому по переходу через ноль нельзя судить о появлении максимума

С помощью компараторов и устройств счёта можно определить длительность сигнала, а по длительности высчитать максимум. Сигнал имеет экспоненциальную зависимость, соответственно увеличение длительности сигналов с увеличением амплитуды будет носить логарифмический характер, поэтому нужна высочайшая скорость счёта, но на данных скоростях современные устройства счёта не работают.

Пиковые детекторы являются отличным решением в такой ситуации. Однако они требуют дополнительных схем для разрядки детектирующего конденсатора.

Идею же перехода производной сигнала через нулевую отметку можно воплотить не только с помощью дифференциаторов. Если исходный сигнал замедлить на десятки наносекунд, а затем взять разность между исходным и замедленным сигналами, то мы получим приращение сигнала за короткий промежуток времени. По изменению знака этой разности можно диагностировать достижение максимума. Замедлить сигнал можно простейшим повторителем на операционном усилителе, для вычитания можно воспользоваться схемой дифференциального усилителя.[9]

Описание структуры устройства

Так как ФЭУ является источником тока, а не напряжения, то в схеме необходимо использовать преобразователь «Ток-напряжение».

Структура аналоговой части представлена на рис. 2.

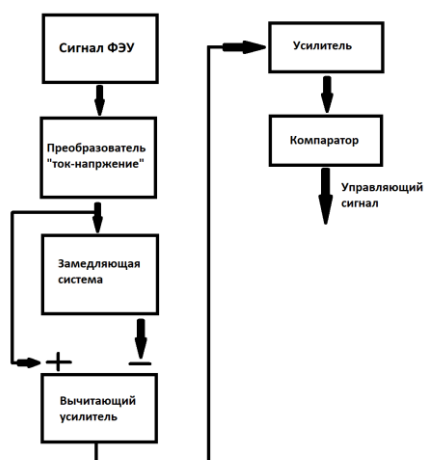


Рис. 2. Структура устройства

Управляющий сигнал, создаваемый компаратором, передним фронтом запускает ячейку памяти, по заднему фронту сигнала АЦП оцифровывает текущее значение ячейки, которое равно максимуму импульса с точностью до утечки заряда конденсатора за время оцифровки.

На Рис. 3. Моделирование работы устройства представлена модель устройства, а именно приёмный тракт устройства обработки аналогового сигнала, ячейка выборки и хранения, позволяющая сохранять значение определённое время и блок аналого-цифрового преобразователя, осуществляющий низкочастотную оцифровку сигнала.

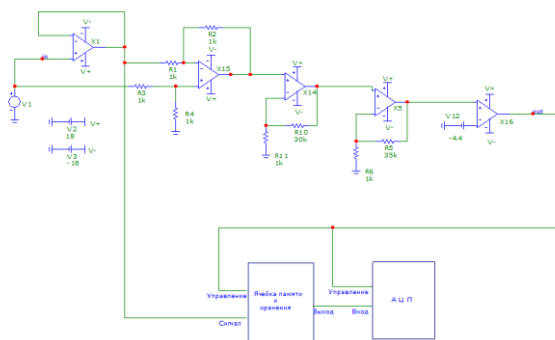


Рис. 3. Моделирование работы устройства

Как можно видеть на рис.4. Графики работы устройства представлены графики, полученные в результате моделирования работы устройства.

Как можно видеть на рис.4. Графики работы устройства представлены графики, полученные в результате моделирования работы устройства.

Рис. компаратор выключается ровно при достижении максимума импульса. Паразитные же импульсы имеют очень малую длительность, за время которой конденсатор в ячейке памяти не успеет разрядиться.

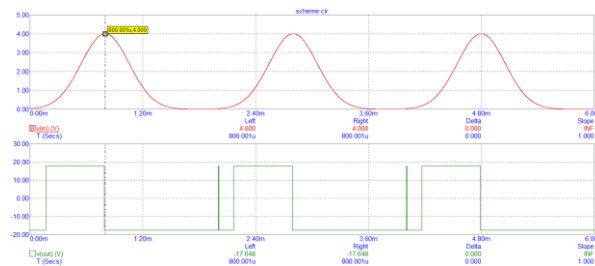


Рис. 4. Графики работы устройства

На первом графике представлен входной сигнал, получаемый с выхода ФЭУ. На втором графике представлена работа компаратора, по срабатыванию которого запускается АЦП микроконтроллера и происходит управление ячейкой выборки-хранения, в которую сохраняется актуальное пиковое значение. Данный подход позволяет отследить конкретный пик и запускать процесс оцифровки только в случае обнаружения пика.

Заключение

В результате работы была получена работоспособная модель устройства являющегося аналоговой частью спектрометрического детектора, регистрирующего ионизирующие излучения. В ходе работы по данной модели была собрана реальная схема, работа которой совпадает с ожидаемыми результатами. Тестирования проводились при помощи лабораторного источника сигналов специальной формы. Устройство будет применено для завершения разработки портативного, автономного спектрометра ионизирующих излучений.

Список литературы

1. Сцинтилляционные счётчики. — URL: <https://www.booksite.ru/fulltext/1/001/008/108/076.htm> (дата обращения 15.12.2022)
2. Introduction to silicon photomultipliers (SiPMs). — URL: https://www.first-sen-sor.com/cms/upload/appnotes/AN_SiPM_Introduction_E.pdf (дата обращения 15.12.2022)
3. Viitanen P. RADIOCHEMICAL METHODS | Radiotracers // Encyclopedia of Analytical Science: Second Edition. Elsevier, 2005. P. 50–58.
4. Legalley D.P. A Type of Geiger-Müller Counter Suitable for the Measurement of Diffracted Mo K X-Rays // Review of Scientific Instruments. American Institute of PhysicsAIP, 2004. Vol. 6, № 9. P. 279.
5. Gavrilyuk Y.M. et al. An ion pulse ionization chamber for spectrometric measurements of low surface α activities // Instruments and Experimental Techniques. 2009. Vol. 52, № 2. P. 173–182.
6. Спектрозональный рентгеновский детектор [Электронный ресурс]. URL: <https://www.elibrary.ru/item.asp?id=37957101> (дата обращения: 12.03.2023).
7. Engelmann E. et al. Tip Avalanche Photodiode -- A new generation Silicon Photomultiplier based on non-planar technology. 2020
8. Dvornikov O. et al. Experimental Studies of the

Silicon Photomultiplier Readout Electronics Based on the Array Chip MH2XA030.

9. Титце, У. Полупроводниковая схемотехника / У. Титце, К. Шенк. — 12-е изд. — Москва : ДМК Пресс, [б. г.]. — Том 1 — 2009. — 832 с. — ISBN 978-5-94120-200-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/915> (дата обращения: 15.12.2022). — Режим доступа: для авториз. пользователей.

РАЗРАБОТКА КОНЦЕПЦИИ USB-НАКОПИТЕЛЯ С ИНДИКАТОРОМ ТЕХНИЧЕСКОГО СОСТОЯНИЯ

Минеев А.В.², Мкртчян Г.А.²,
Насыров Д.Д.², Озорнина М.М.¹

Национальный исследовательский университет
«Высшая школа экономики»,

¹ департамент компьютерной инженерии,

² департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Разрабатываемое устройство позволяет пользователю гарантированно получить оповещения о техническом состоянии устройства различными средствами. Вследствие этого повышается надежность хранения данных на носителе, предоставляется возможность прогнозирования его скорого выхода из строя вследствие изнашивания и старения ячеек памяти, и прочих эксплуатационных факторов, и возможность оповещения об этом. Кроме того, в заявленном устройстве отсутствуют элементы питания и факт отображения информации на информационном дисплее заявляемого устройства не зависит от наличия питания, что обеспечивает постоянную готовность устройства к выполнению его функций по обеспечению сохранности информации и своевременному предупреждению пользователя устройства.

Введение

Деградация элементов памяти в USB-накопителях неизбежна, что приводит к неприятным ситуациям, когда устройство работает некорректно или не работает вовсе. К сожалению, не всегда можно определить проблему и ее причины. В целях решения данной проблемы ведется разработка и создание устройства, представляющего собой USB-накопитель с E-ink дисплеем для отображения основной информации о статусе ресурса и состоянии устройства.

В данной работе рассматривается создание USB-Flash устройства, как с аппаратной, так и программной точки зрения, а также обозреваются логические принципы работы. Кроме того, проводится анализ рынка и сравниваются аналоги.

Аппаратное решение

Разрабатываемое устройство должно состоять из корпуса, внутрь которого помещена электронная плата, содержащая в себе разъем USB типа A,

модуль электропитания, состоящий из регулятора напряжения и кварцевого резонатора Clock PLL, контроллера USB RAM Buffer, микроконтроллера, контроллера памяти, E-ink дисплея, светового индикатора и флэш-памяти NAND type [1]. Это можно увидеть на рисунке 1.

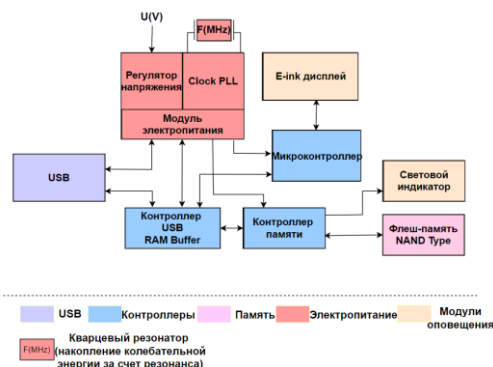


Рис. 1. Функциональная схема

Устройство посредством разъёма USB типа A подключается к компьютеру, через который в модуле электропитания возникает напряжение тока, регулятор напряжения преобразует входное напряжение под необходимые выходные значения, в кварцевом резонаторе Clock PLL накапливается колебательная энергия за счет резонанса. Модуль электропитания сообщает энергию контроллеры (USB RAM Buffer, контроллер памяти, микроконтроллер). Также через USB разъем поступает информационный сигнал на Контроллер USB RAM Buffer (оперативная память), в который передается информация для временного хранения данных при вводе и выводе и идет дальше к микроконтроллеру и контроллеру памяти. Микроконтроллер обеспечивает индикацию режимов работы контроллера и отображает информацию через E-ink дисплей. В свою очередь контроллер памяти передает от оперативной памяти обработанную информацию на флэш-память и световой индикатор. Световой индикатор выполняет роль системы оповещения о состоянии устройства. Во флэш-памяти осуществляется хранение информации посредством логической функции NAND.

Логические принципы работы устройства.

На рисунке 2 можно увидеть алгоритм[2][3].

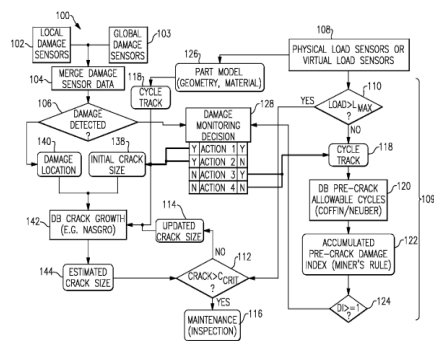


Рис. 2. Функциональный алгоритм

Пользователь соединяет интерфейсный разъём USB 1 устройства с соответствующим разъёмом на ведущем устройстве, которым может быть персональный компьютер или другой совместимый аппарат. В случае нормальной работоспособности ведущего устройства, на интерфейсный разъём USB подаётся питающее напряжение, которое стабилизируется с помощью модуля электропитания до уровня, необходимого для работы микроконтроллера 6 и контроллера интерфейса USB. Микроконтроллер кратковременно блокирует обмен информацией между ведущим устройством и контроллером интерфейса USB и осуществляет расчёт и оценку показателей надёжности.

В его энергонезависимой памяти хранится суммарная статистика работы устройства: количество свободной и занятой памяти, время работы, количество циклов перезаписи, данные о рабочей температуре. На основании этих данных осуществляется математический расчёт вероятности безотказной работы (ВБР) согласно модели надёжности с экспоненциальным распределением. Для расчёта ВБР применяется следующая формула (1):

$$P(t) = e^{-\lambda_{\text{сус}} t_{\text{работы}}}, \quad (1)$$

где $t_{\text{работы}}$ – суммарное время работы устройства, ч;

$\lambda_{\text{сус}}$ – интенсивность отказов флеш-накопителя в циклическом режиме, 1/ч.

При этом $\lambda_{\text{сус}}$ рассчитывается по формуле (2):

$$\lambda_{\text{сус}} = \left(A_1 B_1 + \frac{A_2 B_1}{\pi Q} \right), \quad (2)$$

где πQ – коэффициент качества.

A_1 и A_2 – динамические коэффициенты, зависящие от количества обращений к памяти устройства.

Коэффициент B_1 рассчитывается относительно объёма флеш-накопителя, а также его рабочей температуры, по следующей формуле(3):

$$B_1 = \left(\frac{B}{64000} \right)^{0.25} * \exp \left(\frac{-0.12}{8.63 \cdot 10^{-5}} * \left(\frac{1}{T_j + 273} - \frac{1}{303} \right) \right), \quad (3)$$

где B – объём памяти, бит;

T_j – температурный коэффициент, рассчитываемый по формуле(4):

$$T_j = T_c + \theta_{jc} * P, \quad (4)$$

где T_c – температура работы, °C;

θ_{jc} – тепловое сопротивление «переход-корпус» °C/Вт;

P – максимальная мощность рассеивания, Вт.

Указанная формула частично известна из справочника надёжности MIL-HDBK-217F (Министерство Обороны США, 1991г.). Флеш-память, изготовленная различными производителями, имеет различные характеристики, которые необходимо уточнять непосредственно у производителей.

Далее рассчитывается техническое состояние (Т.С.) устройства по формуле (5):

$$Т.С. = \left(\frac{Q_{\text{в.}}}{Q_{\text{пр.}}} \right) * 100\%, \quad (5)$$

где $Q_{\text{в.}}$ – количество выполненных циклов перезаписи памяти; $Q_{\text{пр.}}$ – предельное число циклов перезаписи памяти. Современные модели флеш-накопителей поддерживают стандарт NAND. Он позволяет перезаписывать данные от 10 000 до 100000 раз.

Затем рассчитывается (по формуле 6) прогнозируемый ресурс (Т), определяющий оставшееся время работы устройства:

$$T = \frac{1}{\lambda_{\text{сус}}}, \quad (6)$$

После расчёта вероятности безотказной работы (ВБР), микроконтроллер анализирует рассчитанное численное значение, принимает решение об оставшемся запасе надёжности устройства и выводит соответствующее сообщение на информационный E-Ink дисплей, которым он управляет, а также, в случае критически низкого запаса надёжности, активирует устройство оповещения 8, которое подаёт звуковой сигнал, оповещающий пользователя, что вследствие износа устройства хранения данных (флеш-памяти), хранить данные на нём более небезопасно. На дисплей выводятся все рассчитанные показатели надёжности: ВБР, техническое состояние, прогнозируемый ресурс. Также микроконтроллер 6 управляет цветом свечения светодиодного RGB индикатора. Когда техническое состояние модуля памяти выше 51%, индикатор мигает зелёным цветом. Если Т.С. ниже 50%, но выше 25%, то светодиод мигает жёлтым цветом. И если Т.С. ниже 25%, то цвет индикатора меняется на красный.

Единожды отображённая на E-Ink дисплее информация сохраняется на нём в течение длительного времени без необходимости поддержания электропитания. Поэтому микроконтроллер отключает питание дисплея и разрешает контроллеру интерфейса обмен данными с ведущим устройством, вследствие чего устройство начинает выполнять функции обыкновенного флеш-накопителя. При этом микроконтроллер анализирует данные, передаваемые по интерфейсу, определяет, произошла ли перезапись ячейки памяти, и суммирует в своей энергонезависимой памяти количество перезаписей и время работы устройства [4][5].

Программное решение

В состав устройства входит микроконтроллер. Микроконтроллер снабжён постоянным машиночитаемым носителем данных с программным обеспечением, сконфигурированным с возможностью реализации функций подсчета произведенного количества циклов перезаписи ячеек памяти, сравнения расчетного значения количества циклов перезаписи ячеек памяти с предельно допустимым количеством циклов перезаписи ячеек памяти, передачи соответствующей информации на информационный дисплей и светодиодный индикатор, с которыми взаимодействует микроконтроллер, или активации устройства звукового оповещения. Носитель данных содержит исполняемые команды, которые при выполнении их устройством приводят к выполнению операций, включающих в себя: регистрацию команды обработки данных, выполняемой контроллером интерфейса USB, определение факта перезаписи ячейки памяти, расчёт остаточной способности ячеек памяти к перезаписи по предельному количеству циклов перезаписи ячеек памяти, заданному производителем модулей памяти, объём занятой памяти и суммарному времени работы устройства, осуществление аналитического расчета и оценки показателей надёжности, расчёт вероятности безотказной работы по экспоненциальному закону распределения с учетом суммарной статистики работы устройства, анализ рассчитанного численного значения, принятие решения об оставшемся запасе надежности устройства и передачу команды на вывод соответствующего сообщения.

Функционал

Основная функция USB-flash накопителя - уведомление пользователя о состоянии устройства. Это будет производиться различными способами: световой индикацией, звуковой индикацией и текстовой и/или графической информацией. Самая важная информация на дисплее - остаточный ресурс накопителя, который отображается при помощи параметров: надежность (показывает процент целостности памяти устройства), а также остаточный ресурс (оставшееся количество часов работы накопителя до выведения его из строя). Кроме того, заявленное устройство будет осуществлять весь функционал стандартного флеш-накопителя.

Аналоги

Прямых аналогов USB-flash устройства с E-ink дисплеем нет, чего нельзя сказать о косвенных.

Проведенный поверхностный анализ рынка флеш-накопителей с интерактивным и информативным дисплеем показал, что на данный момент рынок находится на начальной стадии своего развития.

Хорошим примером является концепт компании Lexar JumpDrive Mercury (Рисунок 3) - USB-флэшка с дисплеем из электронной бумаги, что позволяет, не подключая к компьютеру, проверить состояние памяти флеш-накопителя.



Рис. 3. Устройство Lexar

Также компания “LEVEL UP” изготавливает флеш-карты с LED экраном, на который выводится логотип или другая символика. Изображение на флешке появляется, когда накопитель подключается в гнездо USB. USB-flash устройство компании “LEVEL UP” устройство можно увидеть на рисунке 4.



Рис. 4. Устройство LEVEL UP

Из проведенного анализа видно, что рынок имеет большие перспективы своего развития, но поскольку модели являются концептами или косвенными аналогами, необходимо выделить больше средств и сил на развитие данного направления.

Заключение

В статье были рассмотрены аппаратная и программная часть решения, то есть было рассмотрено как создать устройство с инженерной и программной точек зрения. Помимо этого, рассмотрен логический принцип работы USB-flash устройства, заявлен функционал, а также проведен анализ аналогов на рынке.

Список литературы

1. Flash-память и другие современные носители информации / Кузьмин А. В. - Москва: Горячая линия-Телеком, 2005. - 80 с.
2. ГОСТ 19.701-90. Схемы алгоритмов, программ, данных и систем. Условные обозначения и правила выполнения. 1990 г.
3. Patent US № 2010/0161244 A1, Jun. 24, 2010.

4. Fast contraband detection in large capacity disk drives / Philip Penrose, William J. Buchanan, Richard Macfarlane // Digital Investigation - 2015 - vol.12 - С. 22-29.
5. MIL-HDBK-217F: Reliability prediction of electronic equipment. US Department of defense – 28 February 1995.

ГЕНЕРАТОР АНАЛОГОВОГО СПЛАЙН-СИГНАЛА НА ОСНОВЕ ПЕРВОГО И ВТОРОГО ИНТЕРПОЛЯТОРА НЬЮТОНА

Васьков М.О.

*Московский авиационный институт
(национальный исследовательский университет),
институт №4 «радиоэлектроника,
инфокоммуникации и информационная
безопасность»*

Аннотация

В работе рассматривается возможность использования первого и второго интерполятора Ньютона в устройствах для генерирования сигналов. Данное исследование проводится в рамках разработки модели генератора аналоговых сплайн-сигналов заданной формы на основе интерполяторов таблично заданных функций.

Введение

В настоящее время невозможно представить разработку радиоэлектронных компонентов без использования генераторов радиосигналов заданной формы. Радиоустройства, формирующие сигнал, играют важную роль в контроле, диагностике неисправностей и наладке аппаратуры.

Обычно генераторы оформлены в виде автономных блоков. Их размеры и стоимость ограничивают применение во встроженных устройствах радиотехнических систем. Такое применение генераторов заставляет искать возможности снижения габаритов. В статье рассматривается создание генераторов на основе интерполяторов таблично заданных функций [1], которые сочетают возможности высококачественного синтеза с возможностями улучшения конструктивных характеристик.

Далее рассматривается применение первого и второго интерполятора Ньютона в создании генератора аналогового сплайн-сигнала, представленного в работе [1].

Математический аппарат

Для разработки генераторов заданной формы выбраны интерполяторы таблично заданных функций [2].

Пусть $Y(t)$ - аналоговый сигнал с периодом повторения T . Разделим период T на равные интервалы длительностью τ . В момент времени $t = 0, \tau, 2\tau, 3\tau, \dots, T$ будем фиксировать узловые значения функции и записывать их в память. Устройство считывает из памяти напряжения в узлах функции Y_0, Y_1, \dots, Y_n . Генерирование заданного аналогового

сигнала выполняется аппроксимацией узловых напряжений с помощью интерполяционных полиномов для случая равностоящих узлов.

Используя зафиксированные в памяти узловые значения функции $Y(t)$, генератор передаёт на выход непрерывное значение сигнала $Y_0(t)$, максимально приближенные к периодическому сигналу $Y(t)$.

Для максимального сближения функции воспользуемся полиномами Ньютона.

Первая интерполяционная формула Ньютона:

$$Y_0(t) = y_0 + \frac{\Delta y_0}{1!\tau}(t - t_0) + \frac{\Delta^2 y_0}{2!\tau^2}(t - t_0)^2 + \dots + \frac{\Delta^n y_0}{n!\tau^n}(t - t_0)^n \quad (1)$$

Вторая интерполяционная формула Ньютона:

$$Y_0(t) = y_n + \frac{\Delta y_{n-1}}{1!\tau}(t - t_0) + \dots + \frac{\Delta^n y_0}{n!\tau^n}(t - t_n)(t - t_{n-1}) \quad (2)$$

Для дальнейшего практического использования запишем полиномы (1) и (2) в преобразованном виде, вводя обозначения

$$\alpha = \frac{t - t_0}{\tau} \quad (3)$$

$$\alpha_1 = \frac{t - t_1}{\tau} \quad (4)$$

При этом, ограничиваясь лишь тремя соседними узлами, можно получить достаточно точную модель функции [3].

Используя данное утверждение подставим выражения (3) и (4) в формулы (1) и (2) и запишем в виде зависимости масштабированного сигнала $U(t)$ (коэффициент масштабирования $1/12$) от уровней узловых напряжений u_0, u_1, u_2 на входе синтезатора [1]:

$$U(t) = \frac{1}{12}u_0 - \frac{1}{12}\int_0^a(u_0 - u_1)dx + \frac{1}{12}\int_0^a\left(\frac{1}{2}u_0 - u_1 + \frac{1}{2}u_2\right)dx \quad (5)$$

$$U(t) = \frac{1}{12}(u_0 - u_1 + u_2) + \frac{1}{2}\int_0^{a_1}\left(-\frac{1}{4}u_0 + \frac{1}{3}u_1 - \frac{1}{12}u_2\right)dx + \frac{1}{2}\int_0^{a_1}\left(\frac{1}{12}u_0 - \frac{1}{6}u_1 + \frac{1}{12}u_2\right)dx \quad (6)$$

Аппаратная реализация

При аппаратной реализации сплайн-генератора первый интерполятор Ньютона целесообразно использовать на начальном отрезке функции, а второй - на конечном.

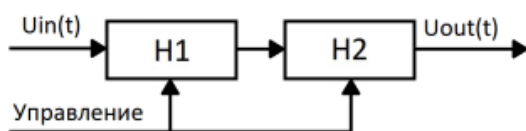


Рис. 1. Схема сплайн-генератора сигнала заданной функции

На рис. 1 показана блок-схема сплайнового генератора, содержащая первый и второй интерполяторы Ньютона ((H1) и (H2)), управляющее устройство.

В аппаратной реализации параболического интерполятора Ньютона (рис. 2) используются математические модели (5) (верхняя секция с выходом Out(1)) и (6) (нижняя секция с выходом Out(2)).

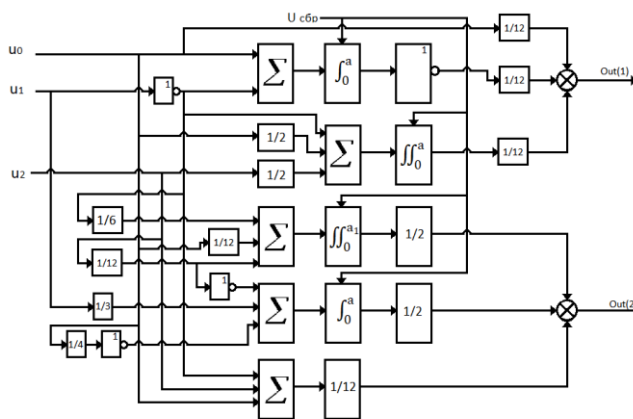


Рис. 2. Параболический интерполятор по формулам Ньютона первого и второго рода

Схема выполнена полностью на аналоговых элементах и содержит 13 резистивных делителя напряжений, семь сумматоров, четыре интегратора и четыре инвертора.

Рассмотрим работу устройства. В ПЗУ записаны узловые значения напряжений синтезируемой функции $f(t)$. Автомат выбора триад узловых напряжений на каждом шаге интерполяции считывает из ПЗУ по три значения, которые передаются на вход синтезатора сигнала (параболического интерполятора Ньютона). После окончания синтеза очередного сплайна автомат сбрасывает интеграторы. Этот процесс циклический и выполняется непрерывно несколько раз до подачи на вход секции блокирующего синхросигнала с частотой $1/T$.

Структура интерполятора (рис.2) является модульной и при необходимости порядок интерполяции может быть изменён путём добавления соответствующих секций сумматора, интегратора и делителей напряжений.

Примеры проектирования

Для получения результата работы генератора и расчета невязки была разработана математическая модель, основой которой является формулы (5) и (6) и схемы (рис. 1) и (рис. 2), на языке программирования Python.

На рис. 3 пунктиром представлены узловые напряжения генератора сплайн-сигналов, модулирующего синусоидальную функцию.

За счёт аппроксимации узловых напряжений с помощью интерполяционных полиномов в аппарате рис. 2 на выходе получается плавное непрерывное изменение сигнала, максимально приближенное к значению заданной функции. Выходной сигнал, построенный по входным узловым напряжениям, на рис. 3 показан сплошной линией.

Сравнивая выходной сигнал с функцией

$$y = \sin(2x) \quad (7)$$

можно определить величину невязки. Она равна

$$\Delta_u = -0.0841181 \text{ В} \quad (8)$$

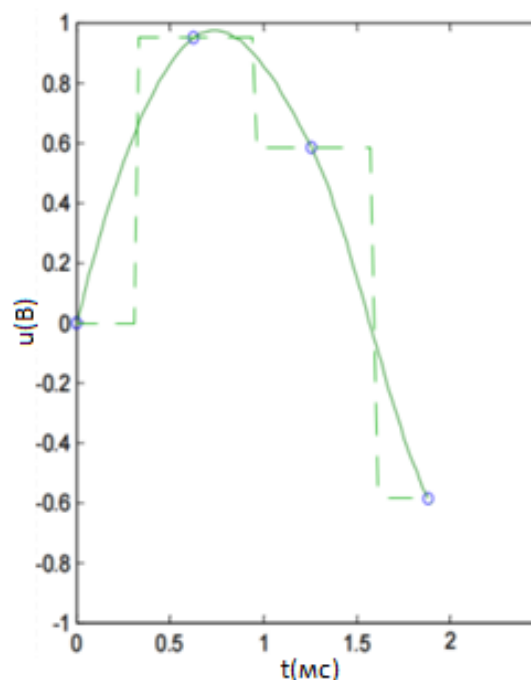


Рис. 3. Пунктиром - входные узловые напряжения, сплошной линией - выходной сигнал

Вывод

Разработана структура параболического интерполятора Ньютона для периодических сигналов заданной формы.

Основным преимуществом предлагаемой технологического решения являются: простота конструкции, малые размеры, модульность, высокая точность, что позволяет использовать его во встроженных устройствах контрольно-диагностических систем.

Список литературы

1. Dembitsky N. L. Spline Signal Generators for a Given Form on Continual Processors. Publisher: IEEE, 2021 Systems of Signals Generating and Pro-

cessing in the Field of on Board Communications, Conference Proceedings, 2021, 9416026.

2. NEWTON'S BACKWARD DIFFERENCE FORMULA // math.iitm.ac.in URL: https://math.iitm.ac.in/public_html/sryedida/caimna/in terpolation/nbdf.html (дата обращения: 10.11.2022).

3. Интерполяция, аппроксимация, оптимизация: анализ и синтез сложных систем: моногр. / А.М. Данилов, И.А. Гарькина. – Пенза: ПГУАС, 2014. – 168 с

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ВЕРИФИКАЦИИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ ПО РАДИАЦИОННОЙ ЭЛЕКТРОПРОВОДНОСТИ ПОЛИМЕРОВ

*Толстиков С.Ю., Муллахметов И.Р.
Национальный исследовательский университет
«Высшая школа экономики»,
аспирантская школа по техническим наукам*

Аннотация

В работе рассматривается алгоритм и результат работы программного обеспечения для сбора и верификации экспериментальных данных испытаний полимерных материалов (далее – ПМ) при условиях, близких к условиям околоземной плазмы.

Введение

Перед запуском космических аппаратов (далее – КА) проводят предварительные расчеты электростатического равновесия всей конструкции, а также напряженностей электрических полей в диэлектриках, возникающих вследствие ионизирующего излучения космического пространства. В этой области широко известны программы NASCAP (для программы NASA Charging Analyzer Program) и ESADDC (ESA Deep Dielectric Charging), DICTAT (Dielectric Internal Charging Threat Assessment Tool) [1], а также разработка МИЭМ [2]. С каждым годом появляются новые ПМ и подбор безопасных значений по заряджению, порогу пробоя и типу пробоя для таких материалов можно получить с помощью экспериментальных измерений [3, 4, 5]. Наше программное обеспечение предназначено для съема и последующей обработки данных экспериментов по измерению радиационной электропроводности (далее – РЭ) полимеров.

Проведение экспериментов

Данные экспериментов снимаются с установки ЭЛА-50/5 в цифровом виде. Дальнейшая обработка происходит в ручном режиме, что накладывает соответствующие ограничения. Оператор производит манипуляции по очистке данных от шумов и помех, а также сравнивает получившиеся экспериментальных значений РЭ с расчётными значениями по модели Роуза-Фаулера-Вайсберга графическим способом. В данном случае верификация получившихся данных полностью зависит от оператора, количественная оценка отсутствует. Пока исследуется несколько об-

разцов ПМ – это приемлемо, но с увеличением количества экспериментов и частой сменой ПМ, появилась необходимость автоматизировать процесс верификации и сравнения получаемых данных с расчётными моделями.

Программное обеспечение

Блок-схема алгоритма работы разработанного программного обеспечения представлена на рисунке 1. На данный момент создана программа в программном пакете Jupyter Notebook. Работа программы будет показана на примере работы с установкой ЭЛА-50/5 в импульсном режиме. Логика работы программного обеспечения для непрерывного режима работы установки ЭЛА-50/5 не будет отличаться, при начале работы программного обеспечения присутствует выбор режима измерений.

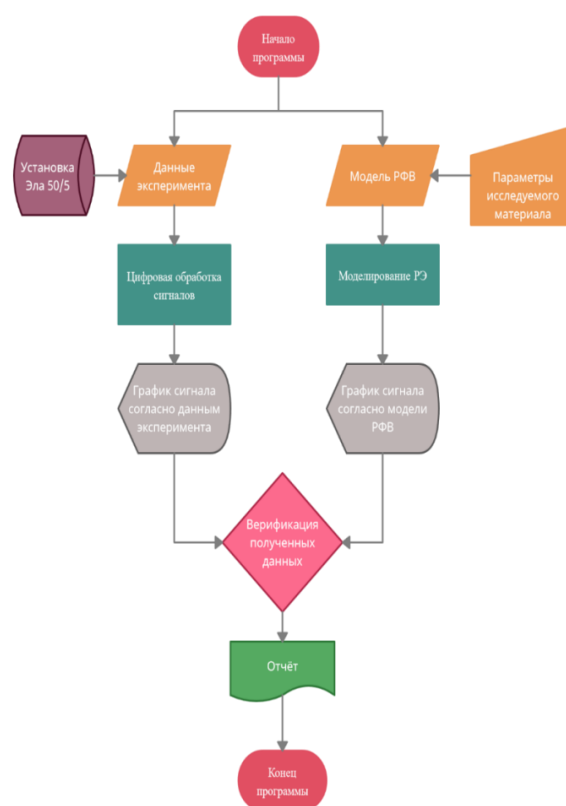


Рис. 1. Блок-схема программного обеспечения по верификации данных РЭ ПМ

Данные измерений поступают в программу напрямую с установки ЭЛА-50/5 или в виде файла с измерениями. На первом этапе обрабатываются данные с измерениями сигнала радиационного тока через образец ПМ, а также измерения сигнала помех. Прототип автоматически строит представленные графики (рис. 2).

Необходимые параметры для расчёта модели РФВ загружаются отдельно в виде подготовленного файла. Пользователю достаточно ввести параметры в интерфейсе программы или выбрать заранее определённый ПМ, параметры выставятся автоматически. В нашем случае, в качестве демонстрации используют-

ся значения изменения РЭ полистирола при температуре жидкого азота. На рисунке 2 видно, что полученные данные совпадают с рассчитанной моделью.

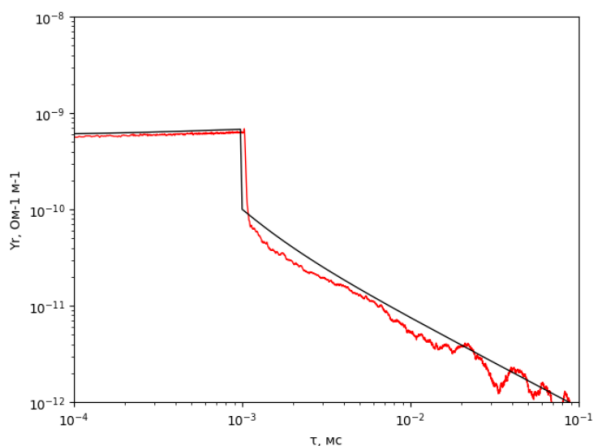


Рис. 2. РЭ полистирола при облучении 1мс импульсом при 79K при моделировании по модели Роуза-Фаулера-Вайсберга (1) и в результате измерения (2)

В данном эксперименте нас интересует оценка аппроксимации. Поэтому для выбранного в программном обеспечении диапазона до 1 мс автоматически будет рассчитана относительная ошибка по формуле для каждой точки, а среднее значение будет описывать среднюю относительную ошибку для данного измерения по формуле (1). Так у исследователя будет понимание насколько экспериментальные данные совпадают с расчётной моделью поведения РЭ ПМ при ионизирующем излучении.

$$N_{\text{мер}} = \frac{1}{n} \sum \left| 1 - \frac{v_{\text{app}}}{v} \right| \quad (1)$$

где $N_{\text{мер}}$ – средняя относительной ошибки измерений, n – количество выбранных совпадающих точек для двух графиков, v_{app} – значение РЭ, полученной во время эксперимента, v – значение по расчётной модели.

Для рисунка 2 удалось выделить порядка 400 точек (рис. 3), при расчёте показывающих, что средняя относительной ошибки, выраженной в процентах, составляет меньше 25%, что хорошо видно на графике (кривые практически совпадают). Если выбрать интересующий нас участок – 200 точек на промежутке от 10^{-4} мс до 10^{-2} мс, то средняя относительной ошибки составляет в процентном выражении менее 20%.

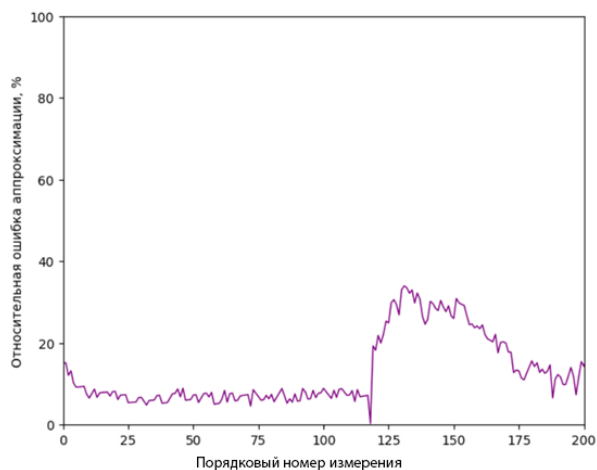


Рис. 3. Относительная ошибка аппроксимации измерений РЭ полистирола при облучении 1мс импульсом при 79K

Заключение

Разработанное программного обеспечения позволяет выбирать интересующий диапазон эксперимента и получать информацию о качестве снимаемых данных в количественном выражении.

Список литературы

1. Rodgers D., Fan L. DICTAT Validation // Spacecraft Charging Technology Conference — Pasadena (USA): SCTC, 2014. — P. 131.
2. Муллахметов И.Р., Толстиков С. Ю., Соболев А. С., Мансырев А.М. Исследование и разработка методов компьютерного моделирования и расчета воздействий электромагнитных импульсов на ракету-носитель в процессе её вывода на расчетную орбиту [Текст] / ВКР по специальности «Компьютерные системы и сети» – М. НИУ ВШЭ, 2020. – 40 с.
3. Paulmier T., Dirassen B., Arnaout M., Payan D., Balcon N., Radiation Induced Conductivity of Space Used Polymers Under High Energy Electron Irradiation // Spacecraft Charging Technology Conference — Pasadena (USA): SCTC, 2014. — P.135.
4. Хатипов С.А., Радиационная электропроводность полимеров, Модель космоса [Текст] / Т.2: Воздействие космической среды на материалы и оборудование космических аппаратов – М.: КДУ, 2007. – С. 361.
5. Brunson J., Dennison JR Hopping Conductivity in Low-Density Polyethylene // Proceedings of the 12th Rocky Mountain NASA Space Grant Consortium NASA Fellowship Symposium. – Salt Lake City, UT, USA, 2007. — 6 p.

Мониторинг потерь радиочастотной энергии

Экспериментальное исследование уровней потерь радиочастотной энергии телекоммуникационного оборудования сотовых операторов осуществлено с помощью измерительной логопериодической антенны КМ6-600/6000 с линейной поляризацией и портативного анализатора спектра с трекинг генератором Arinst SSA-TG R2 [6].

Логопериодическая антенна КМ6-600/6000 предназначена для проведения радиочастотных измерений в диапазоне частот от 600 до 6000 МГц.

Антенна обладает кардиоидной диаграммой направленности (Рис. 3), причем с ростом рабочей частоты возрастает уровень боковых лепестков в вертикальной плоскости с (-21,3) дБ на частоте 1000 МГц, до (-32,3) дБ на частоте 5000 МГц. При этом направленность антенны растет, что определяется уменьшением угла раскрытия по уровню половинной мощности с $145,3^\circ$ до $119,9^\circ$.

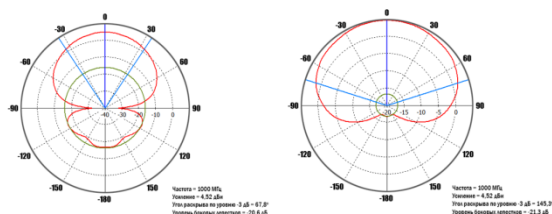


Рис. 3. Диаграммы направленности измерительной антенны КМ6-600/6000 в горизонтальной и вертикальной плоскостях на частоте 1000 МГц

Коэффициент усиления антенны в диапазоне рабочих частот имеет максимум 6,8 дБ (Рис. 4) на частоте 1400 МГц и далее равномерно изменяется от 5,5 до 5,9 дБ на участке 2000 - 6000 МГц.



Рис. 4. Зависимость коэффициента усиления измерительной антенны КМ6-600/6000 от частоты

Для проведения экспериментальных измерений уровней потерь радиочастотной энергии от телекоммуникационного оборудования сотовых операторов в г. Москве были выбраны места, отвечающие основным локациям, которые в течение суток может посещать среднестатистический гражданин. А именно, в помещениях офисного типа, на улицах и центральных площадях, в торговых центрах, метрополитене (как в вагоне поезда, так и на подземных и

наземных станциях), лесопарковой зоне, зоне близ реки, в жилом доме. Диагностирование производилось в старых и новых районах Москвы, в центре и на окраинах, возле промышленных зон. В ходе исследования были получены 26 таблиц со значениями амплитуд в дБ для каждой частоты от 600 до 6000 МГц включительно, далее преобразованных в графики, среди которых 10 были сгенерированы в помещении, а остальные 16 – на открытом пространстве либо в транспорте. Далее собранные данные были сгруппированы по типам локаций:

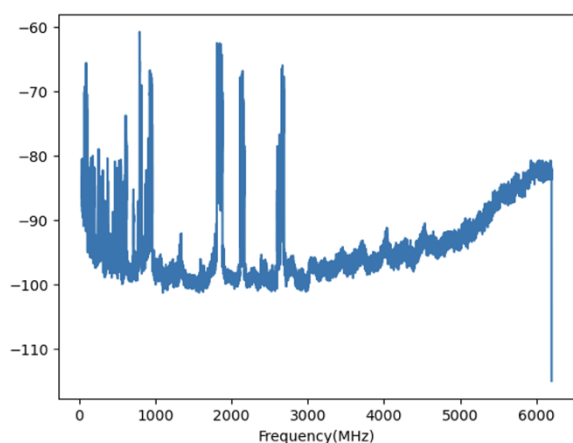
1. Здание МИЭМ НИУ ВШЭ.
2. Улица.
3. Метрополитен.
4. Торговые центры и офисные помещения.
5. Квартира.

Из полученных результатов эксперимента видно, что общая тенденция занятых частот кардинально не зависит от места проведения исследования. На всех измерениях явно отображены пики на частотах 900, 1800 и 2600 МГц. Данные значения экстремумов являются основными частотами базовых станций мобильной связи, которые, несмотря на индивидуальную передачу информации каждому устройству, подвержены сильному загрязнению радиоэфира из-за большой плотности населения и, как следствие, большого числа мобильных устройств, работающих на пиковых частотах. Также хорошо заметно, что амплитуда сигнала базовых станций при измерениях на улице немного выше. Это обусловлено тем, что базовые станции находятся практически в области прямой видимости, и волны не сильно затухают при многочисленных переотражениях, например, от стен домов.

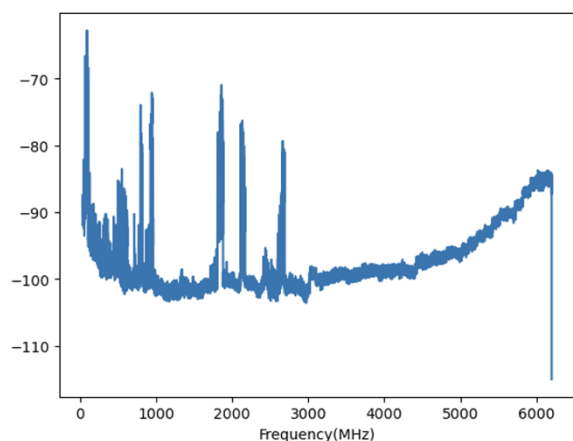
Также следует отметить, что в квартирах одним из наиболее мощных является пик на частоте 2450 МГц, что соответствует сигналам от многочисленных роутеров сети Wi-Fi. Данный максимум практически отсутствует на графике, полученном при измерении на этой частоте, но на улице.

В радиочастотном спектре метрополитена также присутствует небольшой пик от точек доступа Wi-Fi, расположенных в каждом вагоне. Аналогично и для торговых центров (ТЦ) – в них тоже часто присутствует Wi-Fi сеть, причем почти каждый магазин имеет собственную точку доступа, из-за чего амплитуда сигнала существенно выше, чем в квартире. Нельзя не отметить явный пик в самом начале любого из представленных графиков – это пик радиовещательных станций УКВ (FM) диапазона. В городской среде находится избыточное количество различных радиостанций, чем и объясняется крупный всплеск амплитуд сигналов в области от 88 до 108 МГц. Также на всех графиках отлично виден нарастающий пик в 6000 МГц, связанный с функционированием различного измерительного оборудования.

Для дальнейшего анализа используется усреднение экспериментальных результатов по уровню максимальных пиков: оставлено только 2 категории, полученные при исследовании уровней радиочастотных потерь внутри здания и на улице (Рис. 5).



(а)



(б)

Рис. 5. Усредненные зависимости уровней радиочастотных потерь от частоты на улице (а), внутри здания (б)

Заключение

Проведено исследование уровней потерь радиочастотной энергии телекоммуникационного оборудования мобильных операторов сотовой связи с помощью логопериодической антенны (КМ6-600/6000) с линейной поляризацией и портативного анализатора спектра с трекинг генератором Arinst SSA-TG R2. Получены результаты анализа затухания электромагнитных волн диапазонов GSM (UMTS) и Wi-Fi при различных условиях окружающей среды в г. Москве. Из полученных зависимостей следует, что максимальные уровни радиочастотных потерь энергии наблюдаются на частотах 90 МГц, 800 МГц, 1800 МГц, 2100 МГц и 2600 МГц. Данные частоты представляют интерес при создании высокоэффективных интеллектуальных систем сбора радиочастотной энергии для устройств Интернета вещей и киберфизических систем.

Благодарность

Публикация подготовлена в ходе проведения исследования (проект №23-00-003 «Исследование технологий и устройств беспроводного трансфера электромагнитной энергии для высокоскоростных мобильных и нательных устройств Интернета вещей

(IoT/IIoT) и киберфизических систем») в рамках Программы «Научный фонд Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ)» в 2022– 2023 гг.

Список литературы

1. A.A. Yelizarov, I.V. Nazarov, E.A. Zakirova, A.A. Skuridin, D. Rano. Smart Energy Harvesting System for IoT & Cyber Physical Devices / Proceedings of 2022 IEEE Moscow Workshop on Electronic and Networking Technologies (MWENT). IEEE, 2022. P. 1-5.
2. Yelizarov A. A., Nazarov I., Skuridin A., Yaky-menko S., Ikonnikova D. M. Features of Wireless Charging of Mobile and Wearable Devices for the IoT and Cyber Physical Systems/ 2020 International Conference on Engineering Management of Communication and Technology (EMCTECH), 20-22 October 2020, Vienna, Austria. IEEE, 2020. P. 1-4.
3. Y.N.Pchelnikov, A.A. Yelizarov, A.G. Pchelnikov. Wireless Charging Mechanism for Outdoor Appliance // 23rd International Crimean Conference "Microwave Equipment and Telecommunications Technologies". September 8-13, 2013. Sevastopol, Crimea, (KryMiKo-2013). Sevastopol: Weber, 2013. P. 1058-1059.
4. A. A. Yelizarov, A. D. Kuznetsov, I. V. Nazarov and I. R. Lavrukhin, "Prospects for Applications of Small-Sized Antennas on Flexible Substrates," 2020 Systems of Signals Generating and Processing in the Field of on Board Communications, IEEE, 2020, P.1-4.
5. Y.N.Pchelnikov, A.A. Yelizarov, A.G. Pchelnikov. Wireless Charger on Coupled Radial Spirals // Proceedings of the 15th IEEE International Vacuum Electronics Conference (IVEC-2014). Monterey: Institute of Electrical and Electronics Engineers, 2014. P. 473-474.
6. Portable spectrum analyzer with built-in signal generator Arinst SSA TG R2S manual // Arinst.net URL: <https://arinst.net/files/Spectrum-Analyzer-Arinst-SSA-TG-R2s-ENG.pdf> (дата обращения: 20.01.2023).

РЕДУКЦИЯ ЧИСЛЕННЫХ МЕТОДОВ ДЛЯ ПОСТРОЕНИЯ МАКРОМОДЕЛЕЙ ЭКВИВАЛЕНТНЫХ ЭЛЕКТРИЧЕСКИХ СХЕМ КОСМИЧЕСКИХ АППАРАТОВ

Матиян А.А., Вяхирева П.И., Мельникова Е.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии,
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматривается использование различных комбинаций численных методов решения дифференциальных уравнений для построения макромоделей эквивалентных электрических схем космических аппаратов. В ходе выполнения работы рассматриваются численные методы Эйлера и методы Рунге-Кутты различных порядков для поиска наибо-

лее быстродействующего метода с наименьшей погрешностью с целью оптимизации расчетов растекания токов при электризации космических аппаратов.

Введение

В настоящее время существует актуальная проблема электризации обшивки космических аппаратов, которая появляется из-за воздействия на поверхность космического корабля высоко заряженных частиц и космической плазмы. Из-за процесса столкновения частиц с обшивкой может возникнуть электростатический разряд, что может вызывать поломку электрооборудования или полный выход космического аппарата из строя [1]. Чтобы решить данную проблему требуется понимать о динамике распространении тока по кораблю.

Для расчета распространения тока по поверхности космических аппаратов необходимо построить модель редуцированных схем растекания переходных токов по конструкции космических аппаратов [2].

Главной целью работы является исследование и анализ способов редукции дифференциальных уравнений, построение математической модели растекания токов и поиск наилучшей с точки зрения трудоемкости и точности модели, а также программная реализация разработанного метода.

Математическая модель макромоделирования с использованием метода Эйлера

Для начала записывается математическая модель процесса растекания тока:

$$\begin{pmatrix} C_{11} & 0 & 0 \\ 0 & L_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} \dot{\bar{X}}_1(t) \\ \dot{\bar{X}}_2(t) \\ \dot{\bar{X}}_3(t) \end{bmatrix} + \begin{pmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & 0 & G_{23} \\ G_{31} & G_{32} & G_{33} \end{pmatrix} \begin{bmatrix} \bar{X}_1(t) \\ \bar{X}_2(t) \\ \bar{X}_3(t) \end{bmatrix} = \begin{bmatrix} \bar{Y}_1(t) \\ 0 \\ 0 \end{bmatrix} \quad (1)$$

Данные матрицы — симметричные, $a_{C_{13}}, L_{22}, G_{33}$ — диагональные подматрицы

$\underline{X}_1(t)$ — подвектор, включающий в себя напряжения в узлах;

$\underline{X}_2(t)$ — подвектор, включающий в себя токи в узлах;

$\underline{X}_3(t)$ — подвектор напряжений в узлах, к которым не присоединены конденсаторы;

Модель (1) записывается в виде:

$$\begin{aligned} C_{11}\underline{X}_1'(t) + G_{11}\underline{X}_1(t) + G_{13}\underline{X}_2(t) + G_{13}\underline{X}_3(t) &= \underline{Y}_1(t) \\ L_{22}\underline{X}_2'(t) + G_{21}\underline{X}_1(t) + G_{23}\underline{X}_3(t) &= \underline{0} \\ (3)G_{31}\underline{X}_1(t) + G_{32}\underline{X}_2(t) + G_{31}\underline{X}_3(t) &= \underline{0} \end{aligned} \quad (2) \quad (3) \quad (4)$$

Преобразуем уравнение, исключив $\underline{X}_3(t)$. Из (4) получим:

$$\underline{X}_3(t) = -G_{33}^{-1}(G_{31}\underline{X}_1(t) + G_{32}\underline{X}_2(t)) \quad (5)$$

Подстановка $\underline{X}_3(t)$ из (5) в (2) и (3) приводит к следующим выражениям:

$$\begin{aligned} C_{31}\underline{X}_1'(t) + (G_{11} - G_{13}G_{33}^{-1}G_{33})\underline{X}_1(t) + (G_{12} - G_{13}G_{33}^{-1}G_{32})\underline{X}_2(t) &= \underline{Y}_1(t) \\ (6) \quad L_{22}\underline{X}_3'(t) + (G_{21} - G_{23}G_{33}^{-1}G_{31})\underline{X}_1(t) + (G_{21} - G_{23}G_{33}^{-1}G_{32})\underline{X}_2(t) &= \underline{0} \end{aligned} \quad (7)$$

или

$$\begin{pmatrix} C_{11} & 0 \\ 0 & L_{22} \end{pmatrix} \begin{bmatrix} \dot{\bar{X}}_1(t) \\ \dot{\bar{X}}_2(t) \end{bmatrix} + \begin{pmatrix} G_{11} - G_{13}G_{33}^{-1}G_{33} & G_{12} - G_{13}G_{33}^{-1}G_{32} \\ G_{21} - G_{23}G_{33}^{-1}G_{31} & G_{23} - G_{23}G_{33}^{-1}G_{32} \end{pmatrix} \begin{bmatrix} \bar{X}_1(t) \\ \bar{X}_2(t) \end{bmatrix} = \begin{bmatrix} \bar{Y}_1(t) \\ 0 \end{bmatrix} \quad (8)$$

где r - сопротивление резистора в схеме.

Математическую модель можно представить в виде явной системы дифференциальных уравнений:

$$\underline{X}'(t) + A\underline{X}(t) = \underline{Y}(t), \quad (9)$$

$$\text{где } A = \begin{pmatrix} G_{11} - G_{13}G_{33}^{-1}G_{33} & G_{12} - G_{13}G_{33}^{-1}G_{32} \\ G_{21} - G_{23}G_{33}^{-1}G_{31} & G_{23} - G_{23}G_{33}^{-1}G_{32} \end{pmatrix} \begin{pmatrix} C_{11} & 0 \\ 0 & L_{22} \end{pmatrix}^{-1}$$

Первоначальная задача имеет следующий вид:

$$\frac{d}{dt}\underline{X}(t) + A\underline{X}(t) = \underline{Y}(t), \underline{X}(0) = \underline{X}_0 \quad (10)$$

При моделировании процесса (10) с использованием численных методов необходимо Свычислить вектор решения в моменты времени $t_0, t_1, t_2, \dots, t^*, t$. е. $\underline{X}(t_0), \underline{X}(t_1), \dots, \underline{X}(t^*)$.

Давайте рассмотрим одношаговый метод Эйлера. с помощью которого решаются системы уравнений при учете конкретного вида уравнений (10):

$$\underline{X}_{j+1} \approx \underline{X}_j + (-A\underline{X}_j + \underline{Y}_j)h_i \quad (11)$$

Воспользуемся неявным методом Эйлера, учитывая конкретный вид системы уравнений (10). Получим следующую формулу метода Эйлера:

$$(E + h_i A)\underline{X}_{j+1} \approx \underline{X}_j + \underline{Y}_j h_i \quad (12)$$

Преобразуем пункты (11) и (12) для реализации построения новой вычислительной схемы. Выполняем преобразование в некоторую редуцированную модель (макромодель) КА, количество уравнений в которой будет равно $m \ll n$ (второй этап редукции модели). Данные преобразования снижают трудоемкость вычислений на два-три порядка. Для этого запишем оба вида вычислительных процессов в блочной форме:

$$\begin{bmatrix} \underline{U}_1 \\ \underline{U}_2 \end{bmatrix} = \begin{bmatrix} \underline{V}_1 \\ \underline{V}_2 \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \underline{V}_1 \\ \underline{V}_2 \end{bmatrix} h + \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \end{bmatrix} h, \quad (13)$$

$$\begin{bmatrix} E_{11} + hA_{11} & hA_{12} \\ hA_{21} & E_{22} + hA_{22} \end{bmatrix} \begin{bmatrix} \underline{U}_1 \\ \underline{U}_2 \end{bmatrix} = \begin{bmatrix} \underline{V}_1 \\ \underline{V}_2 \end{bmatrix} + \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \end{bmatrix} h, \quad (14)$$

где $[\underline{U}_1 \ \underline{U}_2]^T = \underline{X}_{j+1}^T [\underline{V}_1 \ \underline{V}_2]^T = \underline{X}_j^T$

Пусть подвекторы \underline{U}_2 и \underline{V}_2 содержат по $m \ll n$ искомым коэффициентов решения. На базе выражений (13) и (14) получим редуцированную вычислительную схему. После чего исключим подвекторы \underline{U}_1 и \underline{V}_1 :

$$\begin{aligned} [E_{22} + A_{21}(A_{11}^2)^{-1}A_{12}]\underline{U}_2 &= [A_{21}(A_{11}^2)^{-1}(E_{11} + hA_{11})A_{12} + (E_{22} - hA_{22})\underline{V}_2 + \\ h[-A_{21}(A_{11}^2)^{-1}A_{11}\underline{Y}_1 + \underline{Y}_2] \end{aligned} \quad (15)$$

или

$$\begin{aligned} [E_{22} + A_{21}(A_{11}^2)^{-1}A_{12}]\underline{X}_{2i+1} &= [(A_{21}(A_{11}^2)^{-1}E_{11}A_{12} + \\ + E_{22})\underline{X}_{2i} + h(A_{21}(A_{11}^2)^{-1}A_{11}A_{12} - A_{22})]\underline{X}_{2i} + \\ + h[-A_{21}(A_{11}^2)^{-1}A_{11}\underline{Y}_1 + \underline{Y}_2] \end{aligned} \quad (16)$$

Разработка алгоритмов обработки данных и моделей

В рамках научно-исследовательского проекта мы исследовали описанную выше модель и сделали вывод о ее недостатке: большая трудоемкость построения. В качестве альтернативы предлагаем рассмотреть новую схему на основе методов Рунге-Кутты.

$$X_{n+1} = X_n + hf(t_n + \frac{h}{2}, Y_n + \frac{h}{2} X_n) \quad (17)$$

Для реализации метода Рунге-Кутты 2 порядка (17) потребуется среда разработки. В настоящее время в качестве языка программирования был выбран язык Python 3.7, а средой программирования выбрана PyCharm 2022.3.1. Выбор данной среды программирования основан на существенных преимуществах перед другими вариантами: простая организация проектов, удобный интерфейс для работы с системой поддержки версий git, быстродействие в сравнении с другими средами разработки. Выбор версии языка программирования Python 3.7 основан на поддержке данной версией актуальных для разработки библиотек: scipy, numpy последних версий, а также лучшей оптимизацией по сравнению с предыдущими версиями.

В рамках разрабатываемой программы планируется (см. рис. 1) использование численных методов решения дифференциальных уравнений и алгоритм считывания данных схемы с программ расчета электрических схем, преобразование их к виду матриц, построение схемы и дальнейший расчет схемы различными численными методами.

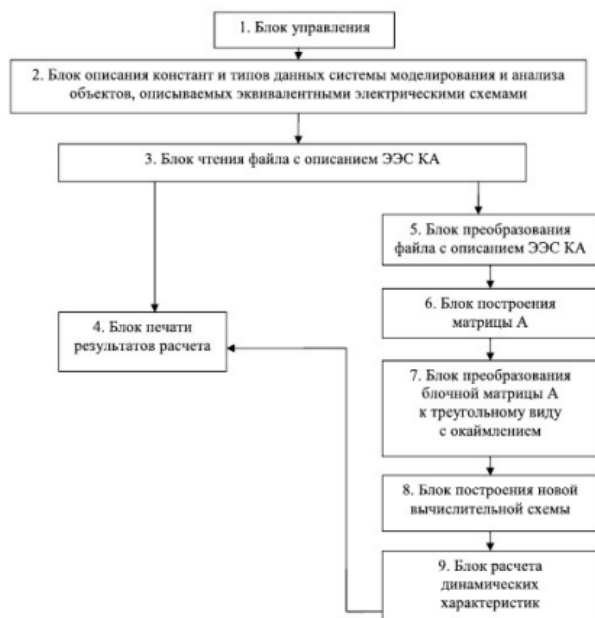


Рис. 1. Иерархическая структура планируемой программы

Далее планируется расчет трудоемкости и погрешности данной модели и использование численных методов Рунге-Кутты разных порядков для по-

иска наилучшей модели по параметрам быстродействия и точности.

Заключение

В результате проведенной работы был проведен анализ способов построения макромоделей, также в работе рассмотрены шаги построения схемы редукции. Недостатком подхода является время построения такой схемы (время сравнимо с трудоемкостью однократного анализа)

Планируется исследование схемы редукции на основе методов Рунге-Кутты 2 и 3 порядка с целью поиска эффективного подхода построения редуцированной схемы.

Список литературы

1. Пожидаев Е.Д. и др. Повышение стойкости космических аппаратов к воздействию поражающих факторов. // «Космонавтика и ракетостроение», 2003, № 1 (30). – С. 32-35..
2. Востриков А. В., Абрамешин А. Е. Вычислительная схема ускоренного метода расчета наводок в бортовой кабельной сети космических аппаратов// Технологии электромагнитной совместимости. 2012. № 3. С. 22-28.
3. Востриков А. В., Абрамешин А. Е., Борисов Н. И. Расчет наводок в бортовой кабельной сети космических аппаратов с помощью макромоделирования на основе методов Эйлера// Технологии электромагнитной совместимости. 2012. № 1 (40). С. 19-24.
4. Востриков А. В., Жадов А. Д., Борисов Н. И., Клышинский Э. С. Разработка и анализ редуцированной вычислительной схемы// Системный администратор. 2014. № 12. С. 90-95.

ВИРТУАЛЬНЫЕ ЛАБОРАТОРНЫЕ СТЕНДЫ ПО КУРСУ ЭЛЕКТРОНИКИ

Ярмиева К.Р.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Разработан программный комплекс для проведения виртуальных лабораторных работ по базовому курсу электроники, позволяющий проводить удаленные лабораторные работы по курсу “Электроника” на платформе LabVIEW, имитируя работу реального лабораторного стенда для задач дистанционного обучения.

Введение

В настоящее время в лаборатории электроники МИЭМ НИУ ВШЭ имеется ряд затруднений при проведении лабораторных занятий по базовому курсу электроники: из-за сильной загруженности лаборатории возникают проблемы при составлении расписания, парк лабораторных приборов устаревает и

снижает свою работоспособность. Также возникают проблемы у студентов, приступающих к выполнению лабораторных работ, так как они недостаточно хорошо знакомы с контрольно-измерительной аппаратурой и методикой выполнения работы. Кроме того, необходимо развитие средств, используемых для курсов дополнительного образования.

Данный проект направлен на решение этих проблем, развивая у обучающихся навыки работы с реальным оборудованием и повышая их готовность к выполнению лабораторных работ по расписанию путем использования средств проекта в самостоятельной работе.

Лабораторные стенды позволяют выполнить все этапы лабораторных работ, обычно выполняющихся в лаборатории, включая вариативность заданий, проверку правильности сборки схемы, работу с измерительными приборами, запись результатов измерений и построение графиков.

Обзор существующих аналогов виртуального стенда

Для подтверждения необходимости разработки стенда был проведен обзор и анализ существующих аналогов.

Первый вариант – удаленный доступ к оборудованию [1], позволяющий работать удаленно и снижая затраты на комплекты реального оборудования. Достоинством является реалистичность поведения всех объектов, но при этом студент не получает навыков работы с оборудованием, так как ограничена интерактивность, также недостатком является необходимость дополнительного ПО и обязательное стабильное подключение к сети Интернет.

Второй вариант – Tinkercad [2]. Данный сервис позволяет самостоятельно собрать схему из виртуальных приборов и провести измерения без использования реального оборудования. Однако недостатком является отсутствие вариантов моделей компонентов и их гибкой настройки, что существенно ограничивает разнообразие выполнения лабораторных работ, а также требование подключения к сети Интернет.

Третий вариант – VR [3], позволяющий наглядно получить навыки работы с виртуальными копиями приборов без реального оборудования. Однако такой вариант требует закупки дорогостоящих VR-устройств и высокой мощности ПК.

Ещё одним вариантом является использование предыдущей версии лабораторного стенда, разработанный ранее в среде LabVIEW [4] и распространяемый в виде exe-приложения (рис. 1). Преимуществом является отсутствие реального оборудования и достаточный уровень интерактивности и наглядности. Также плюсом является возможность гибко настроить как компоненты, так и саму лабораторную работу под нужды обучения. Однако в данном варианте присутствовали недоработки:

- недостаточная сходимость математических моделей полупроводниковых приборов при определенных значениях управляющих напряжений;

- невозможность задействовать произвольное число однотипных контрольно-измерительных приборов и элементов в схеме измерений;

- не универсальная система проверки правильности подключения контрольно-измерительных приборов и объектов измерения, позволяющая проверить только конкретно заданную схему;

- отсутствие универсальных средств ведения протокола измерения в табличном и графическом видах;

- отсутствие универсальной автоматической системы назначения и генерации вариантов лабораторной работы.

На основе обзора было принято решение продолжить работу над виртуальным стендом в среде LabVIEW и устранить указанные недостатки.

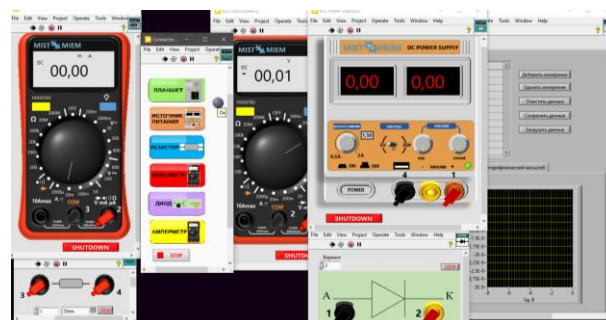


Рис. 1. Внешний вид виртуального стенда на рабочем столе

Описание виртуального стенда

Стенд (рис. 1) состоит из отдельных программ для выполнения следующих лабораторных работ, заключающихся в измерении вольт-амперных характеристик (ВАХ) и определении основных параметров модели исследуемого элемента:

- Изучение статических характеристик диода (стабилитрона) и определение основных параметров его модели для расчёта схем;

- Изучение статических характеристик биполярного транзистора и определение основных параметров его модели для расчёта схем;

- Изучение статических характеристик МОП-транзистора и определение основных параметров его модели для расчёта схем.

Для диода (стабилитрона) это прямая и обратная ветви ВАХ, для биполярного и МОП-транзисторов — входная и выходная ВАХ и сток-затворная и выходная ВАХ соответственно. Схемы измерения представлены на рис. 2.

Стенд состоит из измерительных приборов (вольтметр, амперметр), источника напряжения, резистора и исследуемого элемента. Для исследования транзисторов (биполярного и МОП) предусмотрено по два набора измерительных приборов и два резистора в соответствии со схемой измерения.

Также в составе программы имеется планшет для сохранения результатов измерения в табличном и графическом виде и основное меню со световым индикатором правильности сборки схемы.

Выполнение лабораторных работ заключается в сборке схем, представленных на рис. 2. Соединение элементов происходит с помощью подключения нумерованных проводов к клеммам приборов. Сопротивление резисторов и номер модели исследуемого элемента задаются в соответствии с выданным преподавателем вариантом. При верном подключении световой индикатор не активен, сигнализируя, что можно включить источник питания и начать измерения.

Для этого нужно крутить ручки источника напряжения, задавая напряжение источника. При этом показатели измерительных приборов (вольтметр, амперметр) изменяются соответственно. Данные добавляются пользователем в таблицу на планшете по нажатию соответствующей кнопки, откуда их можно сохранить в формате csv. При необходимости можно загрузить предыдущие результаты и продолжить измерения.

В программе присутствуют всплывающие подсказки при наведении курсора на определенные элементы интерфейса для помощи пользователю в процессе работы.

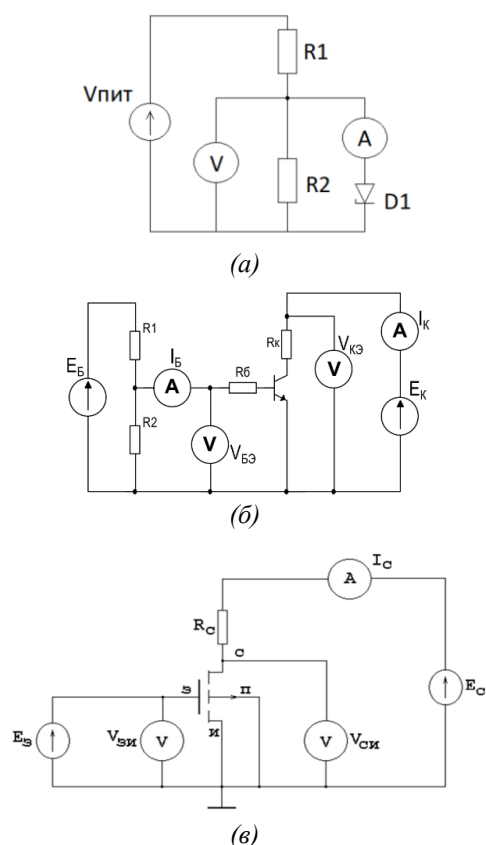


Рис. 2. Схема измерения ВАХ: (а) — диода (стабилитрона), (б) — биполярного транзистора, (в) — диода

Разработка программной части

Для устранения недостатков моделей была разработана система на основе возможностей программы LTSpice [5]. Данная программа позволяет выполнять расчёт параметров собранной схемы, при этом есть возможность проверить правильность подключения пользователем компонентов в логике ошибок. В среде

LabVIEW был разработан модуль convert.vi, обеспечивающий связь основной программы с LTSpice с помощью командной строки. На рисунке 3 представлена часть блок-диаграммы разрабатываемого модуля в среде LabVIEW.

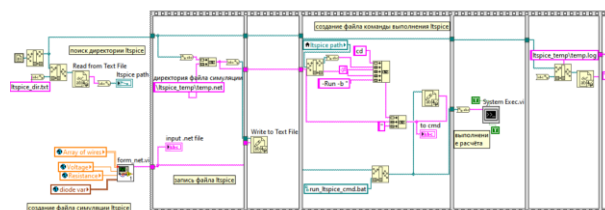


Рис. 3. Часть блок-диаграммы convert.vi

Алгоритм работы: после сборки пользователем схемы в приложении LabVIEW запускается модуль form_net.vi, представленный на рис. 4. Он преобразует данные о схеме в формат, использующийся в LTSpice, с помощью цикла по массиву элементов. В зависимости от типа элемента, соответствующая запись сохраняется в выходной файл. Это текстовый документ в формате .net, содержащий названия компонентов, номера проводов, соединяющих выводы этих компонентов; параметры элементов, такие как напряжение и сопротивление; spice-параметры полупроводниковых элементов (диод, биполярный транзистор, МОП-транзистор); специальные команды для расчёта схемы.

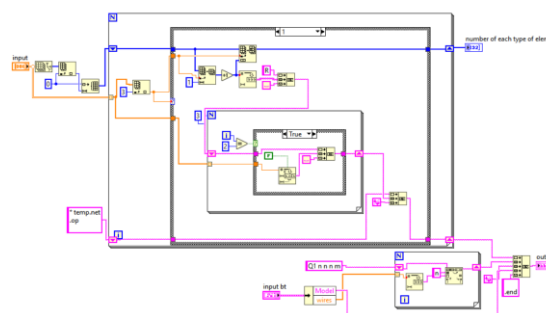


Рис. 4. Часть блок-диаграммы form_net.vi

После формирования файла запускается команда выполнения LTSpice. Для этого используется инструмент LabVIEW System Exec VI, который записывает команду на открытие в LTSpice сформированного net-файла.

Особенностью преобразования является то, что в лабораторных работах используются приборы вольтметры и амперметр. Эти приборы не реализованы в LTSpice, поэтому вместо амперметра при расчёте используется источник с нулевым напряжением, ток на котором измеряется с помощью команды measure, а вместо вольтметра используются узлы соединений, напряжение между которыми также измеряется командой measure [5]. Вид net-файла, открытого в LTSpice, представлен на рисунке 5 на примере биполярного транзистора

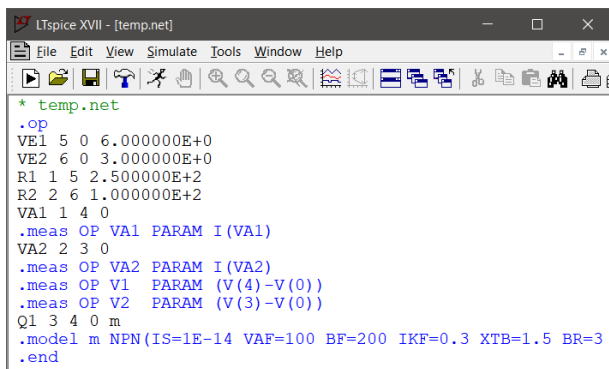


Рис. 5. Вид net-файла, открытого в LTSpice

Файл открывается в LTSpice, при этом происходит расчёт параметров схемы и сохранение результатов моделирования в лог ошибок. Это текстовый файл в формате .log, в котором выполняются заданные команды на измерение, а также записываются ошибки и предупреждения, возникшие при этом. Вид файла приведён на рис. 6.

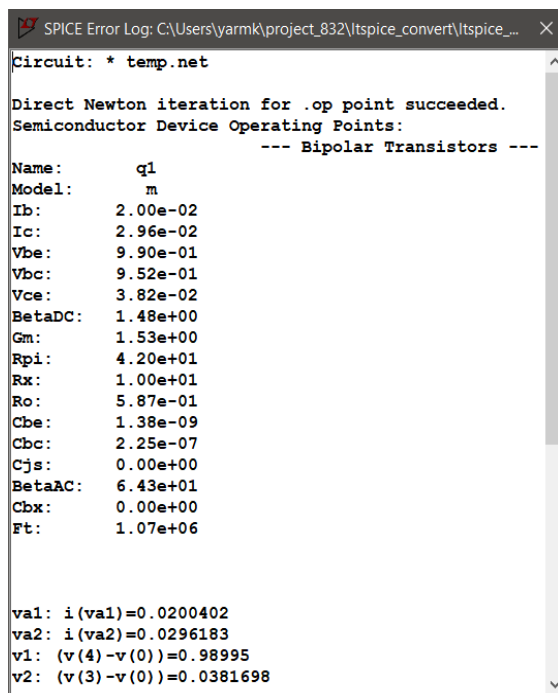


Рис. 6. Вид log-файла, открытого в LTSpice

Данная информация считывается и обрабатывается разработанным модулем. Значения измерительных приборов (вольтметра и амперметра) преобразуются в вид, используемый программой, и передаются к соответствующим приборам. Ошибки, возникшие при моделировании, записываются в лог ошибок в LabVIEW и далее обрабатываются. Это позволяет выполнять проверку правильности сборки схемы пользователем, давая оценку правильности действий пользователя при выполнении лабораторной работы. Все указанные операции выполняются фоновно, пользователь работает только с интерфейсом виртуального стенда (рис. 1).

Описанный модуль решает проблему с ограничениями по диапазону подаваемых напряжений, ранее накладываемый на приборы из-за особенностей моделей, увеличивает гибкость лабораторного стенда, так как позволяет выполнять проверку правильности сборки схемы независимо от вида лабораторной работы, а также позволяет выполнять измерения по схемам вне рамок лабораторных работ по курсу “Электроника”. Также отказ от математической модели компонентов в пользу расчёта в LTSpice позволяет использовать неограниченное количество элементов и приборов в схеме, что также делает виртуальный стенд более универсальным.

Далее планируется улучшить систему генерации вариантов заданий для лабораторных работ и систему проверки правильности выполняемого варианта, а также создать более универсальное средство ведения протокола измерений.

Заключение

В данной работе разработана табличная модель для виртуального воспроизведения реальных схем измерения биполярных транзисторов с внешними сопротивлениями.

Полученные результаты могут быть использованы при проведении лабораторных работ по курсу электроники и в исследованиях характеристик биполярных транзисторов.

Представлены результаты разработки программного комплекса для проведения виртуальных лабораторных работ по базовому курсу “Электроника”, выполненного в среде LabVIEW. Описаны внесенные в виртуальный стенд изменения, расширяющие функционал приложения. Разработанный виртуальный стенд позволяет в рамках дистанционного обучения ознакомиться с функциями и поведением реальных приборов, провести измерения и проанализировать результаты.

Список литературы

1. Евдокимов Ю.К. Дистанционные автоматизированные учебные лаборатории и технологии дистанционного учебного эксперимента в техническом вузе / Евдокимов Ю.К., Кирсанов А.Ю., Салахова А.Ш. // Открытое образование. - 2009. - № 5. - С. 101-115.
2. Муслимова А. З. Использование программно-технических средств в учебном процессе. / Муслимова А. З., Ротман И. М. // Вестник Казахского национального женского педагогического университета. - 2019. - № 4. - С. 119-125.
3. Виртуальные лаборатории : [сайт]. - 2023. - URL: <https://vr-labs.ru/laboratories/electrodynamics/> (дата обращения: 25.01.2023).
4. Тревис Дж. LabVIEW для всех / Джеффри Тревис: Пер. с англ. Клушин Н. А. - М.: ДМК Пресс; ПриборКомплект, 2005. - 544 с.; ил. - ISBN 5-94074-257-2.
5. Володин В. Я. LTspice: компьютерное моделирование электронных схем. / Володин В. Я. - СПб.: БХВ-Петербург, 2010. - 400 с. - ISBN 978-5-9775-0543-7.

ИЗМЕРЕНИЕ ВРЕМЕННОГО РАЗРЕШЕНИЯ РАЗРАБОТАННОГО ДВУХКАНАЛЬНОГО ВРЕМЯ-ЦИФРОВОГО ПРЕОБРАЗОВАТЕЛЯ НА ОСНОВЕ XILINX ZYNQ SOC 7000

Махлуф М.М, Елезов М.С.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В данной работе рассматривается время-цифровой преобразователь (TDC) на плате Xilinx Zynq 7Z007S SoC, который является низкобюджетным ПЛИС. TDC основан на внутренней линии задержки с ответвлениями для точного измерения времени методом старт-стоп. Временное разрешение TDC составляет 17.0 ± 1.3 пс (rms jitter).

Введение

Измерение временных промежутков с высокой точностью между физическими событиями требуется во многих областях науки и промышленности [1-2]. Например, эта методика применяется в лидарах [3], дальномерах, беспилотных летательных аппаратах [4], и робототехнике, усовершенствованных системах помощи водителю (ADAS), системах обнаружения столкновений и других областях. Точность измерений тесно связана с задачами, решаемыми в каждой из этих областей техники. TDC (time-to-digital converter или время-цифровой преобразователь) является основным средством измерения временного интервала.

Существует несколько подходов к реализации конструкции TDC. Большинство из них основаны на интегральных схемах специального назначения (ASIC) и программируемых логических интегральных схемах (ПЛИС). По сравнению с ASIC-TDC [5-6], TDC разработанный на ПЛИС, обладает рядом преимуществ: меньшей стоимостью, высокой гибкостью настроек параметров и меньшим временем проектирования. Поэтому решение с реализацией TDC на ПЛИС становится все более популярным среди разработчиков.

Реализация системы TDC

Блок TDC состоит из грубого счетчика (coarse counter), двух точных счетчиков (fine counter) и энкодера (см. Рис. 1). Грубый (Медленный) счетчик вычисляет количество тактов n_c между стартовым и стоповым импульсами. точный счетчик измеряет количество элементов задержки (T_{Delay}), т. е. разницу во времени между временем срабатывания триггера на стартовый импульс и ближайшим тактовым сигналом clk . Задержка данной линии должна быть несколько больше, чем период тактового сигнала TDC. Используются два точных счетчика, один для стартового импульса, а второй для стопового импульса n_{fstp} (см. Рис. 2).

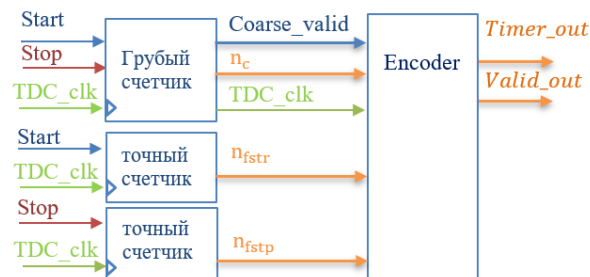


Рис. 1. Структурная схема системы TDC

Следовательно, временная задержка T_{out} между стартовым и стоповым импульсами (1):

$$T_{out} = n_c * clk + (n_{fstr} - n_{fstp}) * T_{Delay} \quad (1)$$

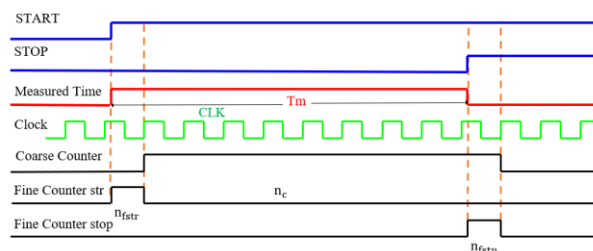


Рис. 2. Схема распространения во времени TDC

Линия задержки

Линия задержки “Time Delay Line” (TDL) представляет собой серию триггеров. Каждый триггер добавляет задержку τ . Зная количество триггеров (см. Рис. 3), можно определить время распространения сигнала t .

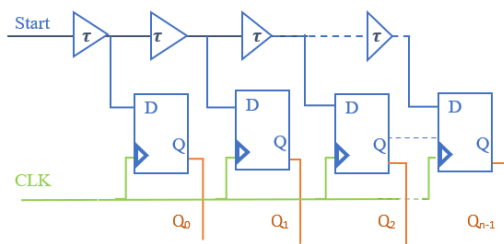


Рис. 3. Концепция субтрактивной линии задержки

Для уменьшения задержки между элементами линии задержки и увеличения точности измерений были использованы линии логики быстрого переноса (англ. fast carry logic) (см. Рис. 4), которые напрямую соединяют SLICE-блоки (группа триггеров) столбца ПЛИСа. Однако, несмотря на такое решение, задержка на каждом шаге имеет разное значение. Поэтому такая схема нуждается в процедуре калибровки.

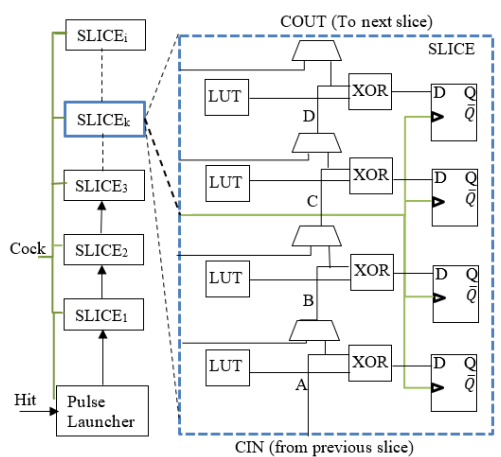


Рис. 4. Xilinx ZYNQ SOC линия логики быстрого переноса (carry4)

Для калибровки был использован метод “statistical code density test” [1]. на входе каждого канала TDC, канале подается периодический сигнал несинхронизированный с тактовой частотой процессора. В результате работы мы должны получить равномерное распределение получения timescode. Однако из-за того, что наша схема неидеальна некоторые timescode будут встречаться чаще, а другие реже. Вероятность встречи timescode определяет время прохождения сигнала по соответствующим элементам линии задержки (см. Рис. 5). Далее, зная тактовую частоту, можно количество событий преобразовать во время прохождения по данному участку линии задержки. Ниже представлена гистограмма временного разрешения (т.е. времени прохождения сигнала по соответствующему участку линии задержки).

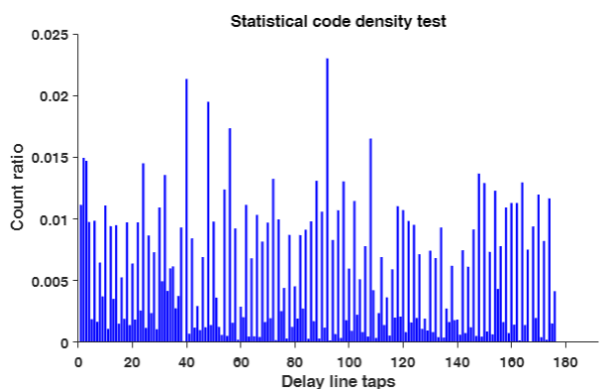


Рис. 5. Распределение отсчетов по элементам линии задержки

Также можно получить калибровочную кривую, которая позволит преобразовать timescode во время задержки сигнала от тактового импульса (см. Рис. 6).

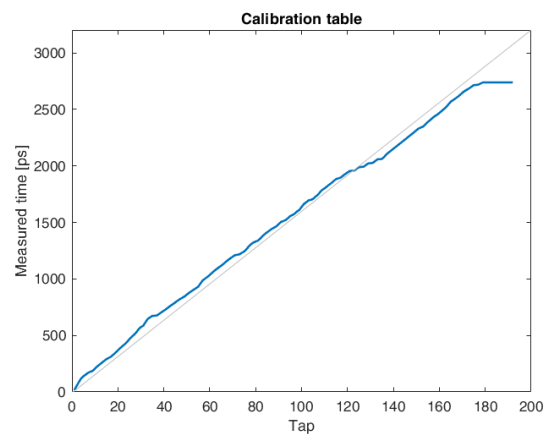


Рис. 6. Калибровочная кривая

Гистограмма измерения постоянного временного интервала

Чтобы определить временное разрешение была построена временная гистограмма распределения времени регистрации стопового импульса относительно стартового импульса (см. Рис. 7). Здесь временное разрешение 17.0 ± 1.3 пс. определяется как среднеквадратичное отклонение от среднего значения времени регистрации стопового импульса.

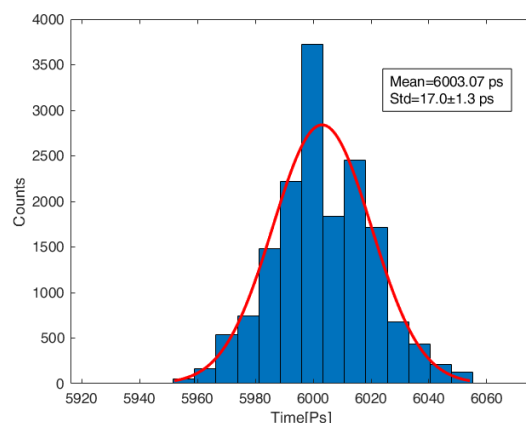


Рис. 7. Гистограмма измерения постоянного временного интервала

Заключение

Было показано, что двухканальный времяцифровой преобразователь может быть реализован на низкобюджетных и легкодоступных ПЛИСах, таких как Xilinx 7Z007S SoC. При этом TDC имеет Временное разрешение 17.0 ± 1.3 пс.

Полученные результаты свидетельствуют о возможности и целесообразности использования подходов TDC для времяпролетных измерений (ToF) на низкобюджетных ПЛИСах.

Список литературы

1. Jozef Kalisz, “Review of methods for time interval measurements with picosecond resolution,” Metrologia, vol. 41, no. 1, pp. 17–32, February 2004.

2. Garzetti, F., Corna, N., Lusardi, N. and Geraci, A., 2021. Time-to-digital converter IP-core for FPGA at state of the art. IEEE Access, 9, pp.85515-85528.

3. Huang, J., Ran, S., Wei, W. and Yu, Q., 2022. Digital Integration of LiDAR System Implemented in a Low-Cost FPGA. Symmetry, 14(6), p.1256.

4. Li, D., Liu, M., Ma, R. and Zhu, Z., 2019. An 8-ch LIDAR Receiver Based on TDC With Multi-Interval Detection and Real-Time $\text{\$ In~ Situ \$}$ Calibration. IEEE Transactions on Instrumentation and Measurement, 69(7), pp.5081-5090.

5. Parsakordasiabi, M., Vornicu, I., Rodríguez-Vázquez, Á. and Carmona-Galán, R., 2021. A low-resources TDC for multi-channel direct ToF readout based on a 28-nm FPGA. Sensors, 21(1), p.308.

6. Burri, S., Bruschini, C. and Charbon, E., 2017. LinoSPAD: a compact linear SPAD camera system with 64 FPGA-based TDC modules for versatile 50 ps resolution time-resolved imaging. Instruments, 1(1), p.6.

РАЗРАБОТКА СТРУКТУРНОЙ СХЕМЫ И АЛГОРИТМОВ ДЛЯ УПРАВЛЕНИЯ ВСТРОЕННЫМИ СРЕДСТВАМИ ЛЕТАТЕЛЬНОГО АППАРАТА ПРИ ПОДГОТОВКЕ К ВЫЛЕТУ ИЛИ ПРОВЕДЕНИЮ РЕГЛАМЕНТНЫХ РАБОТ

Рахимов Р.М.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе показана структурная схема, процесс разработки и принцип работы пульта оперативного контроля в составе комплексной системы управления летательным аппаратом (вертолета).

Введение

Разрабатываемое устройство – пульт оперативного контроля (далее – ПОК) предназначено для применения в составе комплексной системы управления (далее – КСУ), которое будет внедрено в летательное устройство (вертолет) на этапе проведения испытаний. ПОК будет использоваться для наземных работ, связанных с установкой КСУ и ее блоков на борту вертолета, регулировкой, проверкой параметров, поиском неисправностей. Пульт оперативного контроля не предназначен для работы с летным экипажем, работа с ПОК должна осуществляться техника службы наземного обслуживания.

Требования назначения и структурная схема

ПОК должен обеспечивать:

- Отображение на встроенном матричном индикаторе буквенно-цифровой информации о состоянии и значении параметров КСУ;

- Управление встроенной системой контроля КСУ при помощи функциональных кнопок на лицевой панели;

- Двухнаправленный обмен данными с центральным вычислителем управления КСУ.

Функционально ПОК (рис.1) должен включать в себя:

- Системный контроллер на базе однокристального микроконтроллера [1], реализующий протокольную часть периферийных устройств;

- Двустрочный (по 16 символов в каждой строке) матричный индикатор, предназначенный для отображения буквенно-цифровой информации, получаемой от центрального вычислительного устройства (далее – ЦВУ) по кодовой линии связи;

- Шесть кнопок управления режимами встроенных средств контроля на лицевой панели;

- Вторичный источник электропитания, включающий входные фильтры и импульсные DC-DC преобразователи;

- Интерфейс управления встроенными матричными индикаторами, для формирования буквенно-цифровой индикации;

- Интерфейс приёма разовых команд с кнопок управления режимами встроенных средств контроля на лицевой панели;

- Одноканальный мультиплексированный интерфейс с логической организацией по ГОСТ Р 52070-2003 [2] и трансформаторной развязкой, обеспечивающий обмен данными со скоростью 1 Мбит/с по нерезервированному каналу (длина линии связи между ПОК и ЦВУ не более 10 м).

- Одноканальный технологический интерфейс UART (уровня 5 В).



Рис. 1. Структурная схема ПОК

Создание принципиальной электрической схемы и проектирование печатного узла

Принципиальная схема создавалась в САПР OrCad от компании Cadence.

Основные элементы:

- В качестве микроконтроллера для устройства был выбран однокристальный микроконтроллер [1] 1986BE1T от компании Milandr;

- Индикаторы – ИПВ 72A1-4/5X7K [3] - 8 штук (на каждом 4 матрицы 5x7 светодиодов), подключены (рис.2) по протоколу SPI к микроконтроллеру;

- Переключатели (кнопки) – ПКн105-1[4] - 6 штук подключены через схему искрогашения (рис.3) к микроконтроллеру;
- Приёмопередатчик мультиплексного канала (рис.4) по ГОСТ Р 52070-2003 [2] – 5559ИН13У2 [5] – для связи с ЦВУ;
- Первичный преобразователь питания - МДМ7,5-1В05 [6] – вход: 27В, выход: 5В;

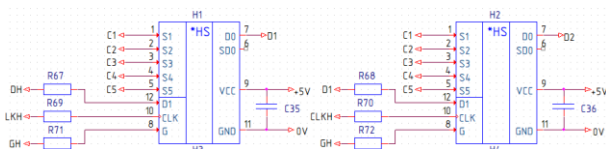


Рис. 2. Часть схемы индикаторов ПОК

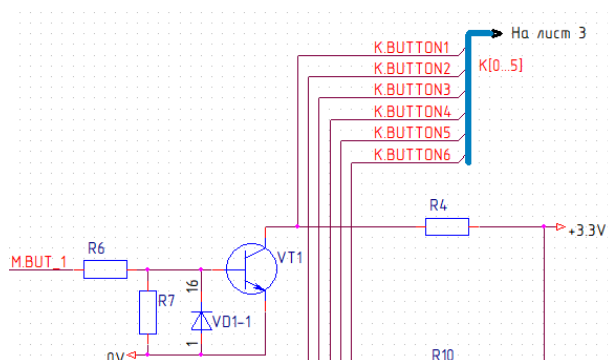
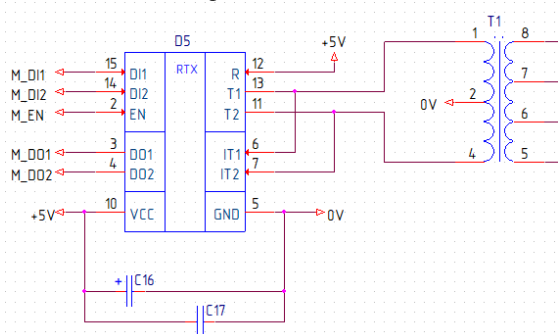


Рис. 3. Часть схемы искрогашения для переключателей



указатель между столбцами (выбор языка, выбор группы параметров, выбор параметра внутри группы, выбор формата отображения параметра) первой строки. С помощью кнопок «плюс» и «минус» оператор увеличивает или уменьшает (либо переключает по списку) значение столбца. После нажатия кнопки «треугольник», ПОК переходит в режим демонстрации параметров (рис.6) в котором убирается курсор с первой строки и активируется вторая строка с названием и значением текущего параметра. В этом режиме теперь можно только переключать параметры (внутри одной группы параметров) с помощью кнопок «плюс» и «минус» и изменять формат отображения параметра с помощью кнопки «вправо». Чтобы вернуться в режим настройки параметров следует нажать кнопку «квадрат».

Заключение

Разрабатываемый пульт оперативного контроля (рис.7) является частью системы КСУ и требуется для организации интерактивного взаимодействия оператора со встроенными средствами контроля комплексной системы управления.

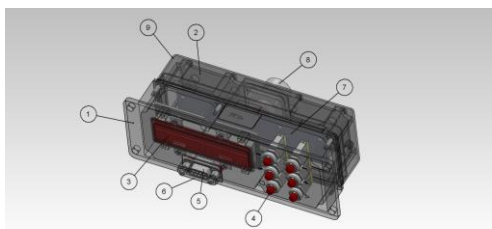


Рис. 7. Внешний вид ПОК

По итогу разработки ПОК будет обладать следующими характеристиками:

- Габариты, мм – 183 x 64 x 90;
- Масса, кг – 1.0;
- Пиковая потребляемая мощность, Вт – 5.

Список литературы

1. Микроконтроллер для авиационной техники 1986BE1T. <https://ldm-systems.ru/f/doc/catalog/LDM-HELPER-K1986BE1QI-FULL/1986%D0%92%D0%951%D0%A2.pdf>
2. ГОСТ Р 52070-2003. Интерфейс магистральный последовательный системы электронных модулей – М.: ИПК Издательство стандартов, 2003
3. Индикаторы ИПА72А1-4/5х7К. <https://diiod.proton-orel.ru/files/indicators/ipv72a1-4-5kh7k-ipv72a-4-5kh7l.pdf>
4. Переключатели кнопочные ПКн105-1В. <https://eandc.ru/pdf/pereklyuchateli/pkn105.pdf>
5. Микросхемы радиационно-стойких приемопередатчиков мультиплексного канала информационного обмена по ГОСТ Р 52070 и MIL-STD-1553В. <https://npofizika.ru/pdf/5559in13.pdf>
6. Низкопрофильные DC/DC модули, серия МДМ 7.5. <https://www.aeip.ru/upload/iblock/bd4/bd445b3f7b93fba11382119b46e0188c.pdf>

ВЛИЯНИЕ ТЕХНИЧЕСКОЙ НАДЕЖНОСТИ ЭЛЕКТРОННЫХ ИЗДЕЛИЙ НА ЭКОНОМИЧЕСКУЮ НАДЕЖНОСТЬ ОРГАНИЗАЦИИ

Старилова Е.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе исследуется взаимосвязь технической надежности электронных изделий и экономической надежности организации. Обоснована актуальность данной работы исходя из необходимости импортозамещения рынка электроники в России. В рамках исследования удалось выявить взаимосвязь гарантии и срока службы электронных изделий на основании данных производителей, что даст право учитывать математически влияние их уровня технической надежности на экономическую надежность организации и исследовать экономическую эффективность России в области электроники.

Введение

В наши дни Россия столкнулась со многими политическими проблемами: санкции, «Специальная военная операция» на Украине и напряженные отношения с некоторыми странами, которые, безусловно, делают вопрос импортозамещения электронных изделий актуальной проблемой. Спрос на продукцию не уменьшился, а изделия должны производиться в достаточном объеме и быть конкурентоспособными. В связи с этим необходимо понять, как надежность изделий рассчитывается в разных странах, сравнить результаты и сделать выводы о том, насколько надежна отечественная продукция по сравнению с зарубежными аналогами. Такое исследование поможет выявить, с какими проблемами могут столкнуться российские компании в условиях импортозамещения и как это повлияет на их экономические показатели.

Цель работы: определить степень взаимосвязи технической надежности электронных изделий и экономической надежности организации, а также оценить последствия импортозамещения ЭКБ в России.

В первой главе рассматривается понятие технической надежности и ее показатели.

Во второй главе исследуется взаимосвязь срока службы и гарантии.

В третьей главе рассматривается понятие экономической надежности и ее показатели.

В четвертой главе описывается взаимосвязь технической и экономической надежности.

В пятой главе рассматривается, каким образом импортозамещение влияет на экономическую надежность.

Техническая надежность изделия и ее показатели

Свойство объекта сохранять в установленных пределах во времени свои параметры и выполнять требуемые функции в заданных режимах называется надежностью. Согласно [1] техническая надежность включает в себя 6 показателей: безотказность, готовность, ремонтпригодность, восстанавливаемость, долговечность, сохраняемость.

Срок службы – это календарная продолжительность эксплуатации объекта от начала эксплуатации до момента достижения объектом предельного состояния. Этот показатель должен быть экономически целесообразным, то есть его эффективное использование должно быть ограничено экономическими последствиями износа объекта. Надежность изделия должна быть достаточной для того, чтобы срок службы изделия был достаточным для его экономической эффективности. Срок службы является характеристикой долговечности.

Наработка на отказ является важным параметром при определении срока изделия. Наработка на отказ – это технический параметр, характеризующий наработку объекта от начала его эксплуатации до отказа. Отношение суммарной наработки восстанавливаемого устройства к количеству отказов, происшедших за суммарную наработку, называется средней наработкой на отказ. Наработка на отказ является характеристикой безотказности, которая необходима для оценки срока службы (долговечности).

Взаимосвязь срока службы и гарантии

Срок службы и гарантия, схожие понятия, однако у них есть ряд отличий. Срок службы может измеряться не только в единицах времени, но и другими единицами измерения (например, км, м и т.д.).

Гарантийный срок – это период, в течение которого производитель обязуется удовлетворить требования потребителя в случае неисправности изделия. Срок службы определяет время, в течение которого изготовитель может предоставить потребителю возможность использования товара.

Взаимосвязь срока службы и гарантии заключается в том, что производитель не может разумно установить гарантийный срок без знания срока службы изделия. Конечно, есть и другие факторы установления гарантии, которые связаны с получением экономической выгоды.

Существует два вида гарантии: та, которую указывает продавец (дополнительная гарантия) и та, которую устанавливает производитель. Технические устройства проходят длинный путь от изготовителя до потребителя, включая транспортировку, таможенное оформление, складское хранение, поэтому второй гарантийный срок больше. Производитель снимает с себя ответственность по гарантийному ремонту, когда гарантия со дня производства истекает.

Обычно продавцы сокращают гарантийные сроки, потому что это выгодно для организации, так как в таком случае ответственность и затраты на ремонт или замену изделия снижаются.

В ходе исследования найдена корреляция между значениями наработки на отказ, которая является неотъемлемой характеристикой при нахождении срока службы, и гарантией, которую дает производитель для 50 блоков питания.

Таблица 1. Наработка на отказ и гарантия для различных блоков питания

№ п/п	Название	Наработка на отказ, ч	Гарантия, лет
1	Cooler Master MWE Bronze 650W V2, 650Вт, 120мм	100000	5
2	Exegate EX292191RUS Серверный БП 900W ExeGate ServerPRO-900ADS	50000	1
3	PHANTEKS REVOLT PRO 1000W Блок питания	100000	5
4	BeQuiet! STRAIGHT POWER 11 850W / ATX 2.4, active PFC	100000	5
...
50	Блок питания 600 Вт PowerMan PM-600ATX-F-BL	60000	1

Коэффициент корреляции r между двумя величинами «Наработка на отказ» и «Гарантия» $r = 0.3754$. Это означает, что связь между величинами слабая, однако она есть. Докажем значимость корреляции.

Найдем случайную величину t (1), которая связана с коэффициентом корреляции r и подчиняется распределению Стьюдента со степенями свободы $Q_{max} - 2$.

Q_{max} – это количество объектов выборки. $Q_{max} = 50$ для частного случая.

$$t = \frac{r \sqrt{Q_{max}-2}}{\sqrt{1-r^2}} \quad (1)$$

Теперь найдем значение $t_{кр}$ (граница критической области – начало значимого интервала) по таблице распределения Стьюдента при заданных $\alpha=0.05$ (объем выборки до 100 единиц) и $Q_{max} - 2$. Если $t > t_{кр}$ или $-t < -t_{кр}$, то t значима – связь выявлена; и наоборот.

Итак, $t = 2.81$ по формуле (1), $t_{кр} = 2.01$. Соблюдается вышеупомянутое условие. Таким образом, связь значима.

Слабость связи объясняется тем, что в выборке были представлены: 1) гарантия продавца, а он обычно устанавливает свой гарантийный срок, не ссылаясь на производителя, который учитывает срок службы изделия; 2) блоки питания разной категории по мощности и стране производителя. Тем не менее это дает право на предположение, что при однородности данных связь будет сильнее.

Экономическая надежность и ее показатели

Способы и методы эффективного взаимодействия всех ресурсов промышленного предприятия, которые выражаются в экономических показателях и дают оценку влияния данного взаимодействия на уровень эффективности функционирования промышленного предприятия во времени для достижения поставленных целей, в совокупности представляют собой механизм экономической надежности промышленного предприятия.

Существуют экономические индикаторы обеспечения эффективности «на входе» и «на выходе».

Обеспеченность процесса технологического развития ресурсами концептуальной модели механизма экономической надежности показывают индикаторы обеспечения «на входе». Эти показатели направлены на персонал, финансы, информацию, поставщиков и инфраструктуру.

Показатели эффективности основных процессов, показатели продукта и показатели удовлетворенности потребителей — результаты функционирования, которые оцениваются на «выходе». К рассматриваемым процессам относятся маркетинг, производство, финансы, организация, функции продукта, надежность продукта, безопасность и дополнительные услуги продукта, деловые связи, клиенты и претензии/рекламации [2].

Учет показателей технической надежности в экономической надежности

Техническая и экономическая надежность тесно связаны. Так, один из рассматриваемых объектов экономической эффективности «на выходе» — надежность продукта, индикаторами которого являются затраты на гарантийный сервис в течение гарантийного срока, количество дефектного продукта. Как уже известно, техническая надежность связана с гарантийным сроком с помощью такого показателя, как срок службы. Что касается дефектов, технически надежное устройство сводит их количество к минимуму, следовательно, экономические показатели организации растут. Чем надежнее устройство, тем меньше дефектов и гарантийных случаев, тем больше сохраненных ресурсов организации.

Импортозамещение и ее влияние на экономическую надежность

Импортозамещение — тип экономической стратегии, направленный на замену импорта товарами, произведенными внутри страны.

Для оценки влияния импортозамещения следует сравнить несколько изделий разных стран. Для простоты исследования, воспользуемся электронным модулем нулевого уровня, а именно классом «Резисторы» группы «Металлодиэлектрический углеродистый» С1-4, т.к. они встречаются в функциональных узлах достаточно часто и составляют более 70% всей его электронной компонентной базы.

Необходимыми данными для расчетной оценки надежности резисторов являются условия приме-

нения. Для примера возьмем условия применения, приведенные в таблице 2 для российского исполнения.

Оценка надежности проводилась на уровне оценки ее показателя — безотказности, в частности, эксплуатационной интенсивности отказов.

Таблица 2. Условия применения для резистора С1-4

Параметр	Численное, значение; ед. измерения
Температура окружающей среды	+30 °С
Коэффициент жесткости условий эксплуатации	1 (офисное помещение)
Коэффициент электрической нагрузки	0.5
Сопротивление	1.8 кОм
Коэффициент приемки	1 (военная приемка)
Время эксплуатации	10 лет

Сводные результаты оценки подобных резисторов, изготавливаемых в разных странах приведены по соответствующим трем справочникам: российскому [3], американскому [4] и китайскому [5] и сведены в таблицу 3.

Таблица 3. Результаты оценки надежности резистора по трем справочникам разных стран

№ п/п	Справочник	Эксплуатационная интенсивность отказов, 1/ч
1	Источник [3]	1.20E-07
2	Источник [4]	6.54E-09
3	Источник [5]	2.2E-09

Итак, сравнивая оцененную эксплуатационную интенсивность отказов с помощью таблицы 3, можно заметить, что российские резисторы имеют меньшую надежность и выходят из строя чаще, чем китайские и американские.

Проблема в том, что наша страна не готова к полному импортозамещению согласно форуму «Микроэлектроника» 2021 и 2022. Объемы производства электронных компонентов в России малы, поэтому создание электронных изделий с российскими компонентами выходит дороже, чем использование зарубежных аналогов. Более того, сроки изготовления компонентов в России гораздо дольше. Чтобы решить эту проблему, предприятия должны быть сами заинтересованы в популяризации своей продукции, быстро подстраиваться под изменяющийся рынок, информировать потребителя, а также наладить масштабов объема производства и его качество.

Заключение

В настоящем исследовании выявлена взаимосвязь технической надежности электронных изделий и экономической надежности организации. В ходе исследования найдена корреляция и ее значимость между гарантией и сроком службы электронных из-

делий. Связь этих показателей доказывает зависимость технической и экономической надежностей на уровне оцениваемых показателей.

Это свидетельствует о том, что для достоверной оценки экономической надежности организации необходима оценка технической надежности электронных изделий, которые она выпускает или использует по назначению согласно условиям применения.

Также, стоит отметить, что в нашей стране объемы производимых ЭКБ малы. Это приводит к тому, что создание полностью отечественных электронных изделий становится невыгодным. Изготовление деталей не только дороже, но и дольше, чем у зарубежных конкурентов.

Данное исследование также помогло выявить, что надежность российской ЭКБ ниже, чем у иностранных (табл. 3), что приведет к частым поломкам и выходам из строя изделий, что, в свою очередь, повлияет на экономические затраты компаний.

Список литературы

1. ГОСТ Р 27.102-2021. Национальный стандарт Российской Федерации. Надежность в технике. Надежность объекта. Термины и определения.: дата введения 01.01.2022. – Москва: Российский институт стандартизации, 2021. – 36 с.
2. Яковлева Е. В., Терентьева А. Р. Экономическая надежность технологического развития // Омский научный вестник. Сер. Общество. История. Современность. 2021. Т. 6, № 1. С. 98–108. DOI: 10.25206/2542-0488-2021-6-1-98-108.
3. Надежность ЭРИ 2006: справочник. М.: МО РФ, 2006. 641 с.
4. Chinese Standard GJB/Z 299C-2006. Reliability prediction handbook for electronic equipment. [Дата введения 2006-01-01]. China. 282 p.
5. Reliability prediction of electronic equipment MIL-HDBK-217F: Military Handbook. AMSC N/A, 1991. 205 p.

ОЦЕНКА НАДЕЖНОСТИ СКОРОСТНОГО ВЫЧИСЛИТЕЛЬНОГО МОДУЛЯ- СИНХРОНИЗАТОРА ОПТИЧЕСКОЙ ЛИНИИ ДЛЯ УПРАВЛЕНИЯ ФАЗОВЫМИ МОДУЛЯТОРАМИ ОПТИЧЕСКОЙ СИСТЕМЫ

Гронин Д.С.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе рассматривается метод оценки надежности скоростного вычислительного модуля, применяемого для синхронизации оптической линии и для управления фазовыми модуляторами. Расчет

надежности выполняется в специализированной программе «Асоника К-СЧ».

Введение

Волоконно-оптическая связь - это метод связи, который передает сигналы в виде света и использует оптическое волокно в качестве среды для передачи этих оптических сигналов из одного места в другое. Сигналы, передаваемые по оптическому волокну, преобразуются из электрических сигналов в свет, а затем снова преобразуются из света в электрические сигналы на приемной стороне. Передаваемые данные могут быть в виде голоса, видео или телеметрических данных и должны передаваться на большие расстояния или в локальной сети. Волоконно-оптическая связь применяется в различных коммуникационных приложениях благодаря своим хорошим результатам в передаче данных на большие расстояния с высокой скоростью [1]. Скоростной вычислительный модуль для синхронизации оптической линии, управляющий фазовыми модуляторами, внедренный в систему квантового распределения ключей, должен отвечать высоким требованиям надежности, ввиду чего проводится данная исследовательская работа.

Факторы, влияющие на оценку надежности

Для оценки надежности первоочередно необходимо понимание какие именно факторы влияют на надежность. Глобально, все факторы делятся на технические, которые в свою очередь подразделяются на конструкционные и производственные, и эксплуатационные, подразделяемые на объективные и субъективные. Конструкционные включают в себя выбор схемных и конструктивных решений, выбор групп и типов комплектующих элементов и материалов, качество отработки схемы и конструкции и качество отработки технологии. Производственные - качество материалов и комплектующих элементов, получаемых от смежников, эффективность входного контроля материалов и комплектующих элементов, отработанность технологического процесса, эффективность контроля качества и уровень организации рекламационной работы и анализа отказов. К объективным факторам относятся режимы работы, уровни внешних воздействующих факторов (таких как климатические, механические, агрессивные среды и другие), уровни воздействия специальных факторов (ионное излучение, электромагнитные излучения различной природы) и качество электропитания. К субъективным соответственно - квалификация обслуживающего персонала, организация хранения и качества ремонта, организация контроля и технического обслуживания и организация сбора данных об отказах.

Не менее важной составляющей являются факторы, обуславливающие отказы ЭРИ (рис.1)

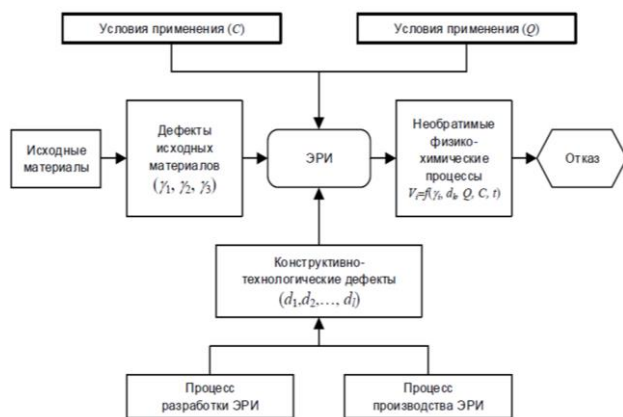


Рис. 1. Факторы обуславливающие отказы ЭРИ

Полный перечень коэффициентов моделей расчета интенсивности отказов ЭРИ и их характеристик приведен в Таблице 1 стр. 9-12 Справочника “Надежность электрорадиоизделий” 2006г. [2].

Ввиду того, что в данной исследовательской работе необходимо определение среднего времени наработки до отказа и вероятность безотказной работы, очевидно, что некоторые элементы или группы элементов будут не соответствовать заявленным требованиям. Для корректности принимаемых решений по повышению надежности, обязательно понимание модели интенсивности отказов (рис. 2).

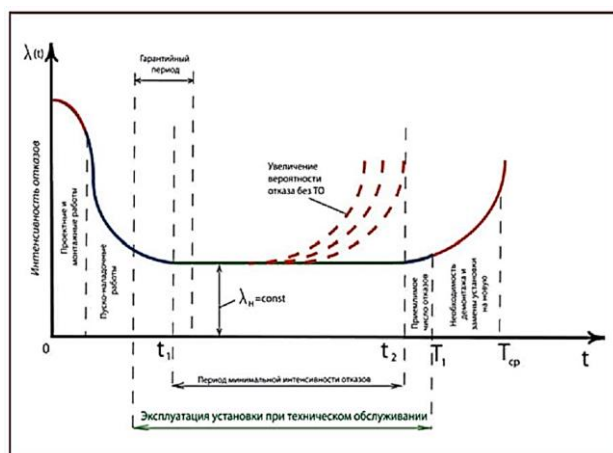


Рис. 2. Интенсивность отказов

Постановка практической задачи

Основной задачей исследования является определение показателей безотказности всех электрорадио изделий (ЭРИ) и внесение изменений в конструкцию с целью достижения необходимой надежности. Для этого в программе Асоника К-СЧ производится расчет среднего времени наработки до отказа, эксплуатационной интенсивности отказов и вероятности безотказной работы для каждого элемента, а также для групп элементов (рис. 3) для определения общей надежности устройства. Цель и ценность работы заключается в достижении требуемых показателей надежности, а именно Время

наработки до отказа должно составлять 8760 часов и вероятность безотказной работы должна быть не ниже 0,98 см. (рис. 4) для скоростного вычислительного модуля (рис. 5).

$\lambda \approx 2,0606579136727E-5$
Интенсивность отказов в режиме ожидания : 2,95011293465297E-7

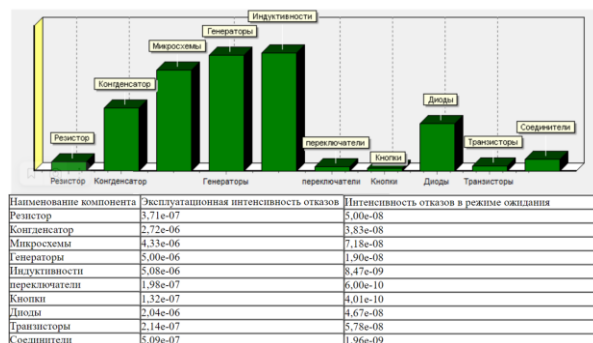
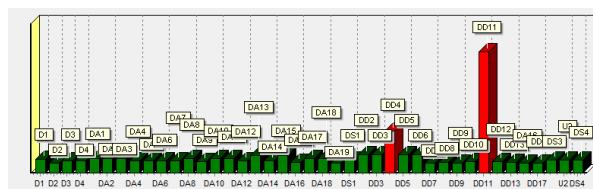


Рис. 3. Группы элементов



```

graph TD
    Start([Начало]) --> 1[/1. Анализ технического задания/]
    1 --> 2[2. Выбор варианта структурной схемы РЭС]
    2 --> 3[/3. Определение требований к выходным характеристикам ФУ и выбор ЭСХ/]
    3 --> 4[4. Расчет электрических характеристик ФУ]
    4 --> 5[/5. Разработка конструкции РЭС/]
    5 --> 6[6. Геометрическое моделирование конструкции РЭС]
    6 --> 7[7. Анализ радиационных полей конструкции РЭС]
    7 --> 8[8. Моделирование тепловых, механических и других режимов работы РЭС]
    8 --> 9[9. Расчет показателей безопасности и долговечности]
    9 --> 10{10. Удовлетворяет требованиям ТЗ?}
    10 -- да --> End([Конец])
    10 -- нет --> 11[/11. Анализ вкладов ЭРИ, параметров и факторов/]
    11 --> 12[12. Внесение изменений в ТЗ]
    12 --> 13[/13. Замена ЭРИ  
Новый вариант ЭСХ  
Резервирование ЭРИ/]
    13 --> 14[14. Внесение изменений в конструкцию]
    14 --> 11
    14 --> End
  
```

Рис. 6. Схема проведения исследования по расчету надежности

Ожидаемые результаты

Ввиду того, что требования к исследуемому модулю завышены по запросу заказчика, придется вносить изменение в конструкцию с применением резервирования. Результатом исследовательской работы будет вариант конфигурации устройства, которое отвечает требованию о безотказной работе в течение двух лет.

Результаты исследования

В результате первичного моделирования был выявлен недостаток надежности у двух классов элементов (рис. 7 и 8). В первичной конфигурации устройство не соответствует заявленным требованиям надежности, ввиду чего, было принято решение произвести расчеты для конфигураций с нагруженным и ненагруженным общим резервированием.

В качестве результатов проверки гипотез представлена графическая аналитика (рис. 9, 10, 11). Ввиду особенностей эксплуатации прибора лучшим решением будет использование нагруженного общего резервирования, так как этот вариант удовле-

творяет двум из трех параметров надежности в реальном рабочем режиме. Однако, ни один из вариантов не обеспечивает необходимого уровня эксплуатационной интенсивности отказов, так как класс «Микросхемы» и «Индуктивности» имеют крайне высокие уровни интенсивности отказов, значительно превышающие $5,86 \cdot 10^{-6}$ [1/ч].

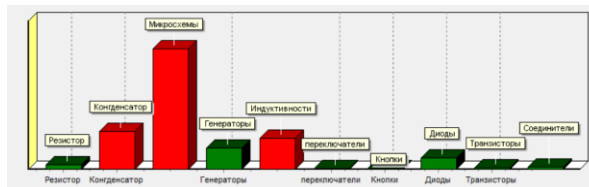


Рис. 7. Уровни эксплуатационной интенсивности отказов для каждого класса ЭРИ в режиме предельных нагрузок при температуре $+60^{\circ}\text{C}$ (класс подсвечивается красным цветом при превышении $5.86\text{E-}6$ [1/ч])

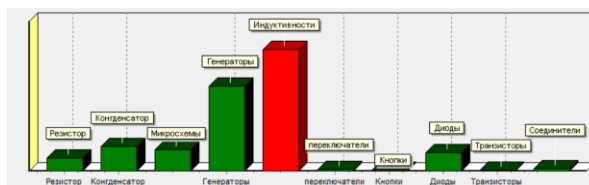


Рис. 8. Уровни эксплуатационной интенсивности отказов для каждого класса ЭРИ в реальном режиме рабочих нагрузок при температуре $+60^{\circ}\text{C}$ (класс подсвечивается красным цветом при превышении $5.86\text{E-}6$ [1/ч])

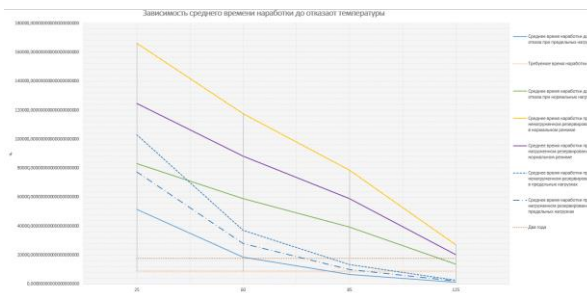


Рис. 9. Зависимость среднего времени наработки до отказа от температуры в разных режимах и конфигурациях

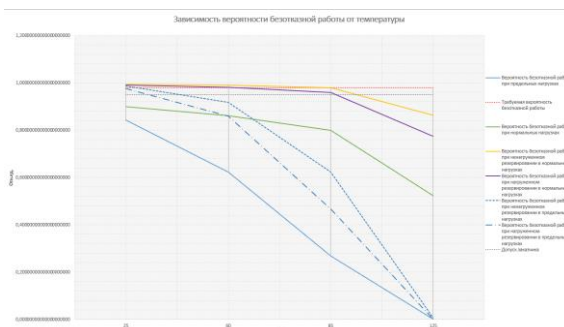


Рис. 10. Зависимость вероятности безотказной работы от температуры в разных режимах и конфигурациях

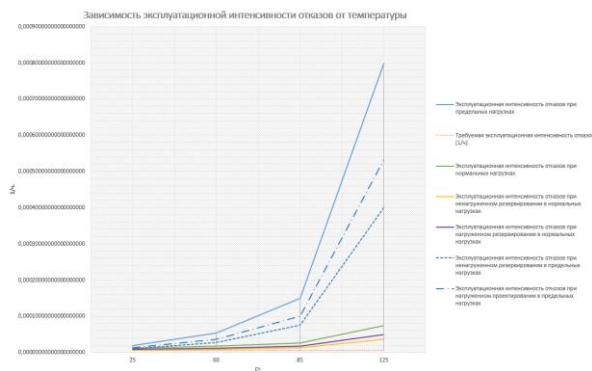


Рис. 11. Зависимость эксплуатационной интенсивности отказов от температуры в разных режимах и конфигурациях

Вывод

По результату первичного моделирования и проверки гипотез были выявлены ЭРИ наиболее подверженные выходу из строя. Учитывая, что применение резервирования не обеспечило должного уровня интенсивности отказов, а температуры, учитываемые при моделировании, являются результатом теплового излучения соседних устройств в стойке, рекомендуется либо изолировать исследуемый прибор путем нанесения теплоотражающего покрытия на кожух, либо заменить интегральные микросхемы и трансформаторы на аналоги с большей номинальной мощностью и более высокими допусками НТД по температуре соответственно.

Список литературы

1. Якушенков П.О. Оптические модуляторы света // PHOTONICS Russ. Technosphera JSC, 2019. Т. 13, № 6. С. 594–602.
2. Надёжность ЭРИ ИП: Справочник. - М.: МО РФ, 2006.
3. С.М. Боровиков, И.Н. Цырельчук, Ф.Д. Троян Расчет показателей надежности радиоэлектронных средств. – Минск : БГУИР, 2010. – 68 .

ИЗМЕРЕНИЕ И МОДЕЛИРОВАНИЕ ВАХ ПОЛЕВЫХ ТРАНЗИСТОРОВ С УПРАВЛЯЮЩИМ PN-ПЕРЕХОДОМ

Горбунов К.А., Исмаил-Заде М.Р.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе представляется получение ВАХ (вольт-амперная характеристика) JFET-транзисторов BF862 и BF545B и получение их SPICE моделей при комнатных температурах. На основе полученных SPICE моделей, проведено моделирование схемы маломощного предусилителя.

Введение

При проектировании устройств различной сложности обороной, космической и бытовой промышленности производителю важно понимать характеристики транзистора. Таким образом ключевым показателем схемотехнического элемента является вольт-амперная характеристика. Данный показатель иллюстрирует характер работы транзистора. Кроме этого, производителю важно иметь возможность моделирования устройства на этапе его проектирования. Для решения данной проблемы используются SPICE-модели[1].

Разрабатываемая SPICE-модель транзисторов, представляет собой совокупность различных параметров полевых транзисторов, важных для схемотехнического моделирования. Для получения модели в данной работе используется VTO, ALPHA, BETA, GAMMA, LABDA параметры[2].

В рамках данной работы в первой части приводится описание процесса измерения вольт-амперных характеристик JFET-транзисторов[3] BF862 и BF545B, а также получение их SPICE-моделей, через подбор параметров в программе схемотехнического моделирования IC-CAP[4]. Во второй части работы приводится результат моделирование работы транзисторов в схеме маломощного предусилителя на одном транзисторе с обратной связью.

Моделирование SPICE-моделей JFET транзисторов

Различают два основных показателя полевых транзисторов это ВАХ для тока стока от напряжения на стоке и для тока стока от напряжения на затворе при фиксированном напряжении затвора в первом случае и стока во втором. Таким образом с помощью специального устройства были сняты показатели с транзистора BF862 и BF545B. На рис. 1-2 приведены ВАХ и транзисторов.

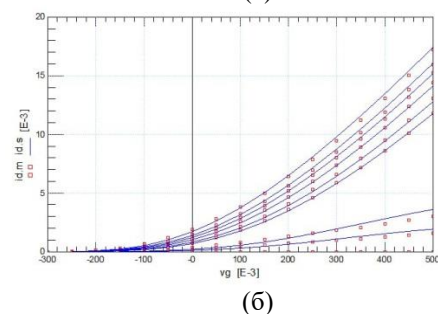
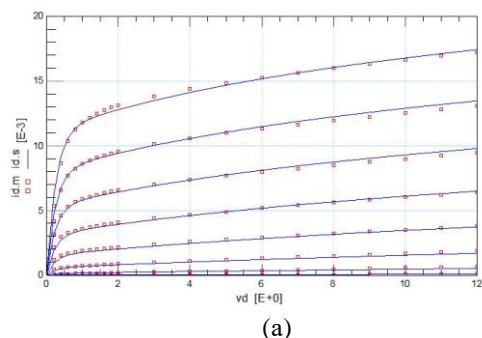


Рис. 1. ВАХ транзистора BF862, (а) характеристика $idvd$, (б) характеристика $idvg$

Следующим этапом является получение SPICE-модели транзисторов. Для данной работы была выбрана модель 3-го уровня (модель Статза), включающая в себя большее количество параметров, что позволяет получить более точный результат. Также данная модель подразумевает работы в температурах от -100 до $+125$ °C, что подходит для данной работы.

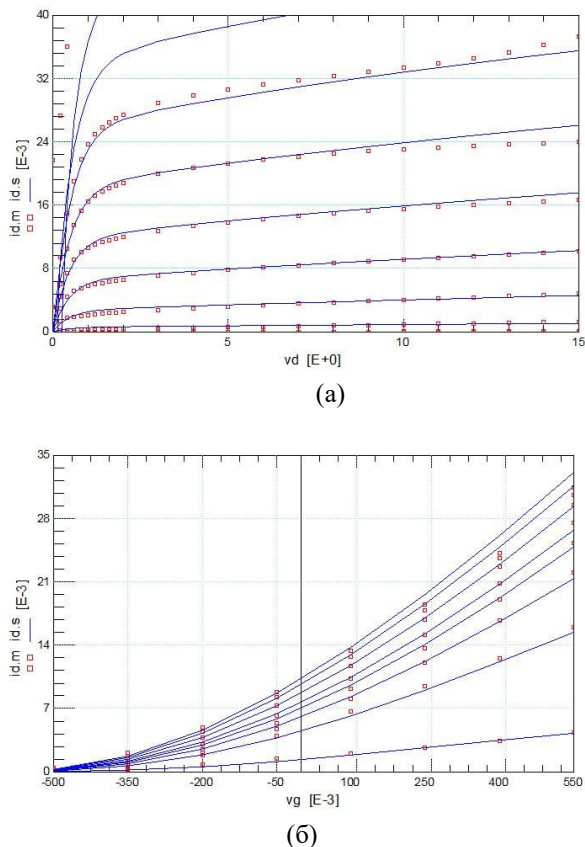


Рис. 2. ВАХ транзистора BF545B, (а) характеристика i_{dvd} , (б) характеристика i_{dv}

Используя программу IC-CAP, был запущен процесс моделирования, во время которого проводится подбор выбранных параметров модели. Отвечающие за пороговое напряжение V_{TO} , крутизну $BETA$, модуляция длины канала $LAMBDA$, коэффициент напряжения $ALPHA$, коэффициент снижения порогового напряжения $GAMMA$. Результаты моделирования также представлены на рис. 1 для BF862 и на рис. 2 для BF545B.

По результатам моделирования получены точное соответствие с ВАХами, измеренными с самих транзисторов.

Моделирование схемы

Следующим этапом работы является моделирование схемы с транзисторами. Для каждого транзистора была проведена модулирование и получены результаты. В качестве результат моделирования представлены АЧХ (амплитудно-частотная характеристика), так как схема используется в виде малошумного предусилителя.

Для работы была выбрана схема малошумного предусилителя с обратной связью (рис. 3). Схема на одном JFET транзисторе, в качестве обратной связи используется резисторы в соединении типа звезда, а также на конденсаторе блокирующий постоянный ток. Входной сигнал подается на затвор, а выходной сигнал снимается со стока.

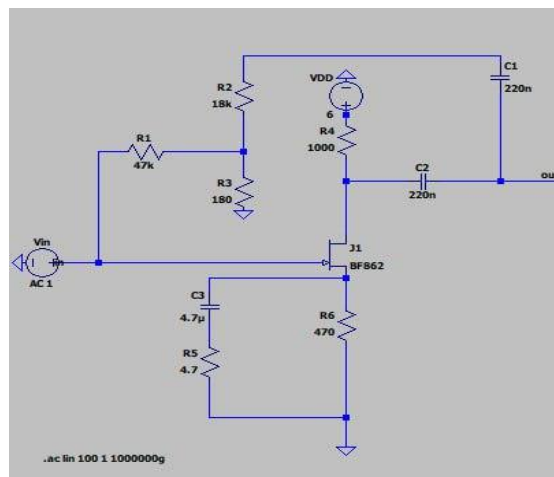


Рис. 3. Моделируемая схема

Моделирование проводится в программе схемотехнического моделирования LTSPICE. SPICE модели передаются в программу и их основе создается новый транзистор. Далее этот транзистор помещается на схему, устанавливаются параметры моделирования. По полученным данным можно определить, рабочий диапазон частот и оценить стабильность схемы.

Заключение

Результаты проведенной работы над исследованием полевых транзисторов с управляемым PN-переходом демонстрируют потенциал использования данных транзисторов в устройствах малошумных, которые выполняют функцию предусилителя.

На основе SPICE модели 3-го уровня определены основные параметры транзистора, которые необходимы для моделирования в схеме или устройстве.

Список литературы

1. Fernández-Martínez P. et al. A new vertical JFET power device for harsh radiation environments //Energies. – 2017. – Vol. 10. – №. 2. – p. 256.
2. Wang L. et al. Analysis and suppression of unwanted turn-on and parasitic oscillation in SiC JFET-based Bi-directional switches //Electronics. – 2018. – Vol. 7. – №. 8. – p. 126.
3. Fernández-Martínez P. et al. A new vertical JFET power device for harsh radiation environments //Energies. – 2017. – Vol. 10. – №. 2. – p. 256.
4. Petrosyants K. O. et al. Automation of parameter extraction procedure for Si JFET SPICE model in the $-200...+110$ °C temperature range //2018 Moscow Workshop on Electronic and Networking Technologies (MWENT). – IEEE, 2018. – p. 1-5.

Принцип работы данного генератора состоит в непосредственном взаимодействии магнитного поля подвижного первого магнитного сердечника 1 и подвижного второго магнитного сердечника 2 с магнитными полями индуцированных ими токов в цилиндрической катушке 5. При встречном движении первого магнитного сердечника 1 и второго магнитного сердечника 2 внутри цилиндрического корпуса 11 в пространстве между одноименными полюсами первого магнитного сердечника 1 и второго магнитного сердечника 2 и охватываемом цилиндрической катушкой 5 с радиальной обмоткой напряженность поля возрастает как в следствии встречного движения первого магнитного сердечника 1 и второго магнитного сердечника 2, так и в следствии индукционной реакции радиальной обмотки цилиндрической катушки 5, в которой индуцированный ток образует магнитный поток, препятствующий движениям пер-

вого магнитного сердечника 1 и второго магнитного сердечника 2. В результате этого нарастание напряженности магнитного потока в рабочем пространстве происходит в прогрессии. Соответственно величина тока в цилиндрической катушке 5 возрастает до максимального значения при минимальном зазоре между полюсами первого магнитного сердечника 1 и второго магнитного сердечника 2.

Широкое применение линейных вибрационных микрогенераторов сдерживается сложностью их проектирования в связи с недостаточной изученностью. Для решения этой проблемы с помощью метода конечных элементов разработаны модели электромагнитных полей линейных микрогенераторов [9, 10] и имитационные модели микрогенераторов в составе АСЭ на их основе (рис. 3).

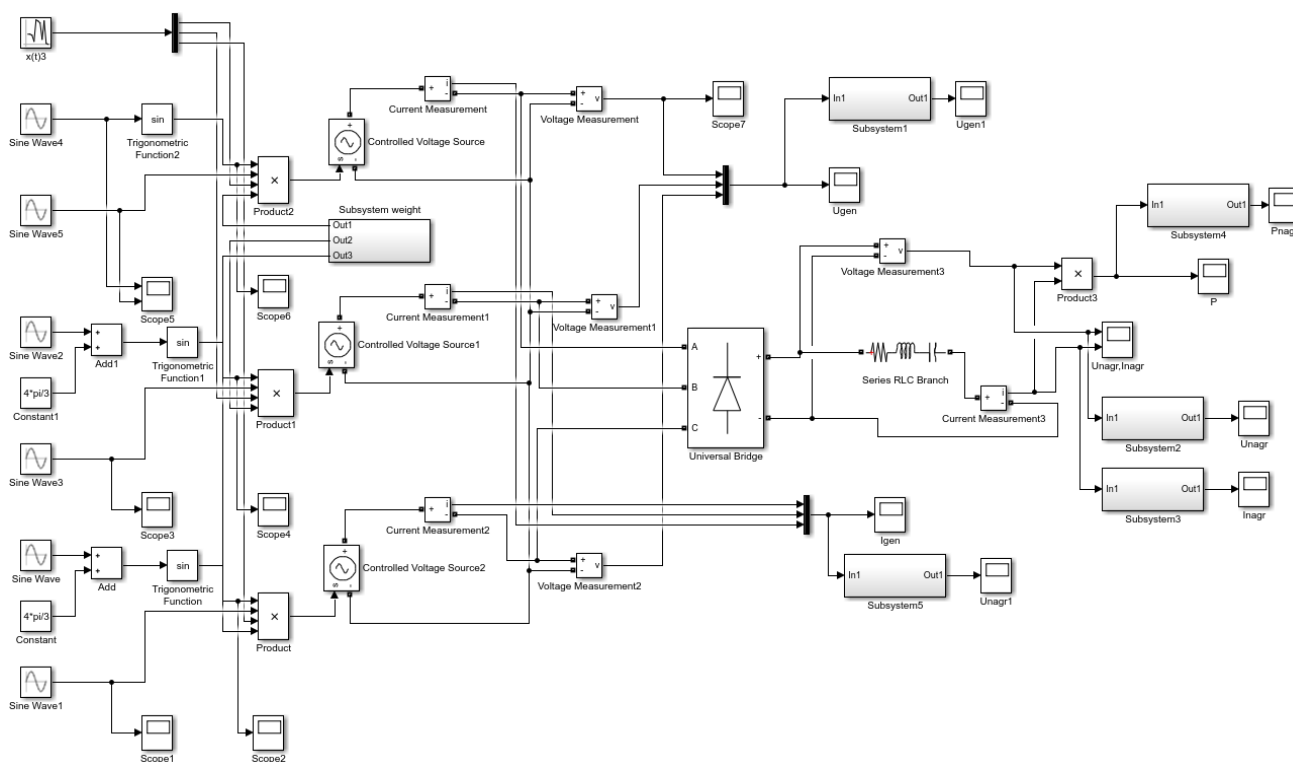


Рис. 3. Имитационная модель трехфазного вибрационного микрогенератора, учитывающая параметры и характер движения индуктора.

Данная имитационная модель позволяет осуществлять анализ динамических процессов, протекающих в микрогенераторе, а также проводить оценку эффективности энергопреобразования в трехфазном синхронном ВМГ с постоянными магнитами. Модель включает в себя силовые блоки, и виртуальные осциллографы для визуализации результатов расчета, блоки измерения электрических величин:

- напряжения,
- электрического тока,
- средней мощности.

Характер выходных параметров микрогенератора (напряжения, тока и мощности) можно определять, задавая массу индуктора, жесткость механических

пружин и (или) демпфера, характер внешних возмущающих колебаний, величину возможной присоединяемой массы (в случае работы микрогенератора на подвижном основании).

В случае, если внешние возмущающие воздействия изменяются по синусоидальному закону, то выходные параметры микрогенератора (ток, напряжение) за 3 периода колебаний имеют вид, представленный на рисунках 4 и 5.

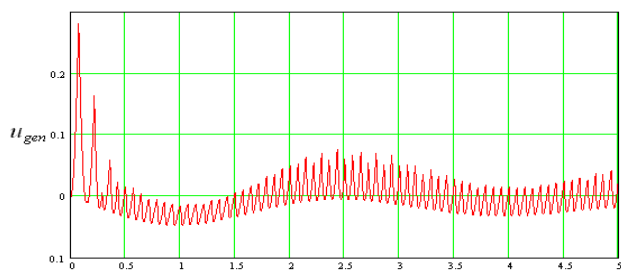


Рис. 4. Характер изменения напряжения U_{gen} в трехфазном синхронном микрогенераторе

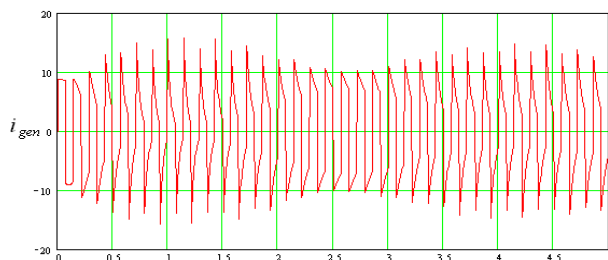


Рис. 5. Характер изменения тока i_{gen} в трехфазном синхронном микрогенераторе

На основе разработанных моделей и алгоритмов проведен анализ динамических процессов и оценка эффективности энергопреобразования в одно- и трехфазных синхронных вибрационных микрогенераторах. Определены основные направления развития и использования разработанных конструкций АСЭ на основе вибрационных микрогенераторов. Разработаны рекомендации для выбора значений геометрических и конструктивных параметров вибрационных микрогенераторов, обеспечивающих повышение эффективности систем автономного питания узлов беспроводных сенсорных сетей.

Заключение

Разработанные модели линейных вибрационных микрогенераторов обеспечивают возможность создавать АСЭ для питания узлов беспроводных сенсорных сетей. Применение АСЭ на основе вибрационных микрогенераторов позволит сделать WSN более надежными и автономными.

Список литературы

1. Молчанов, Д.А., Кучерявый Е.А. Приложение беспроводных сенсорных сетей // Электросвязь. – 2006. – № 6. – С. 20–23.
2. Сергиевский М.В. Беспроводные сенсорные сети // КомпьютерПресс. - №8. - 2007.
3. Шахов В.В., Мигов Д.А., Соколова О.Д. Беспроводные сенсорные сети, оснащенные средствами получения энергии из окружающей среды // Проблемы информатики. – 2014. - №4. – С. 69 – 79.
4. Бочарников И., Уитакер М. Energy Harvesting. Новый этап в развитии автономных устройств // Компоненты и технологии. – 2010. – №8. – С. 146 – 149.
5. Гольцова М. Аккумуляция кинетической энергии из окружающей среды. // Электроника НТБ. – 2011. -№7 - С. 78 - 85.

6. Синютин Е.С. Новые интегральные решения для разработки сборщиков энергии из окружающей среды // Инженерный вестник Дона. – 2014. - №4.

7. Балагуров В.А., Галатеев Ф.Ф., Ларионов А.Н. Электрические машины с постоянными магнитами. - М.: Энергия. - 1964. - 479 с.

8. Высоцкий В.Е., Сеницин А.П. и др. Линейные магнитокоммутационные генераторы для систем электропитания автономных объектов.// Известия высших учебных заведений. Электромеханика. – 2014. - № 5. - С. 26 – 31.

9. Каунг Мьят Хту, Довгаль В.М. К вопросу построения конечно-элементной модели вибрационного микрогенератора // Научные ведомости. Серия: Экономика. Информатика. - 2019. Том 46, № 1. – С. 99 – 107.

10. Каунг Мьят Хту Анализ выходных характеристик вибрационного микрогенератора автономной системы электроснабжения маломощных потребителей в режиме холостого хода // Инновационная наука. – 2019. - №3. - С. 39 – 43.

РАЗРАБОТКА СТРУКТУРЫ СИСТЕМЫ СБОРА И ОБРАБОТКИ ДАННЫХ О ПОВРЕЖДЕНИЯХ ДОРОЖНОГО ПОЛОТНА

Ротай А.Д.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе разработана структура системы сбора и обработки данных о повреждениях дорожного полотна, а также принцип работы такой системы. Проведен обзор имеющихся аналогов системы.

Введение

В России насчитывается более 1.5 млн км автомобильных дорог, среди которых почти миллион – с твердым покрытием. По этим дорогам ежедневно ездят миллионы автомобилей. Естественно, с течением времени дороги приходят в негодность и требуется ремонт или замена дорожного полотна. Скорость ремонта же, в особенности в отдаленных районах или на не самых проходных улицах – оставляет желать лучшего. Существующие способы сбора данных о повреждениях не позволяют охватить все имеющиеся дороги, а следовательно – некоторые выбоины могут «жить» годами.

Решением имеющейся проблемы может стать система, состоящая из устройства, устанавливающегося в каждую новую продаваемую машину в стране, в автоматическом режиме отслеживающее повреждения полотна при помощи комплекса датчиков, веб-сервера, обрабатывающего информацию и веб-интерфейса для пользователей и администраторов.

Задачи: изучить имеющиеся аналоги разрабатываемой системы, разработать структуру и принцип работы системы. В разделах работы проведен анализ источников, рассмотрена разработка структуры устройства и принцип его работы.

Обзор систем для оценки состояния дорожного полотна

В статье «Road Damage Detection using YOLO with Image Tiling about Multi-source Images» Dongjun Jeong, Jua Kim представили нейронную сеть для детектирования повреждений дорожного полотна, обученную на снимках из 6 стран. [1] Данная нейросеть показала F1-меру 0,6744. Но данная система подразумевает использование в оснащенных камерами автомобилях или в беспилотных автомобилях. Данное решение не отвечает задаче массовости (а также дешевизны) разрабатываемой системы.

В статье «The Assay of Potholes and Road Damage Detection» Dr. Yusuf Perwej в соавторстве с Ankit Kumar, Neha Kulshrestha, Ashish Kumar Srivastava и Chandan Mani Tripathi описывают разработку системы для смартфонов с ОС Android для сбора информации о повреждениях дорожного полотна. [2] В дальнейшем изображения обрабатываются нейронной сетью с точностью в 84%. Данная система не отвечает требованию, согласно которому необходим сбор данных в автоматическом режиме, так как в данном случае пользователь самостоятельно делает снимки и загружает их в приложение.

Разработка структуры системы

Разрабатываемая система имеет несколько важных требований:

1. Дешевизна, чтобы не сильно влиять на стоимость автомобиля для конечного пользователя.

2. Устойчивость к вибрационным и ударным нагрузкам.

3. Устойчивость к различным погодным изменениям (водо- и пыленепроницаемой)

4. Отсутствие непосредственного участия пользователя для полноценного функционирования.

Так как система подразумевает детектирование существующих повреждений, устройство, используемое в автомобиле, должно быть оснащено датчиками вибрации для детектирования ударной и вибрационной нагрузки на автомобиль, акселерометрами для детектирования изменений положения колес, GPS-модулем для получения координат повреждения дороги и замера скорости автомобиля, а также LTE-модулем для передачи данных на удаленный сервер для их обработки.

Также система включает в себя удаленный сервер, на котором происходит обработка данных, собираемых автомобилями, их обработка и автоматическое формирование запросов на устранение повреждений дорог в ответственные организации. Данный сегмент системы будет реализован с использованием Node JS, Express, PostgreSQL.

Для администрирования системы будет использоваться пользовательский веб-интерфейс, также адаптированный под мобильные устройства. Разработка пользовательского интерфейса будет вестись с использованием открытых фреймворков и библиотек – React JS как основа структуры, TypeScript для типизации, eCharts для визуализации данных, взаимодействие клиент-сервер будет

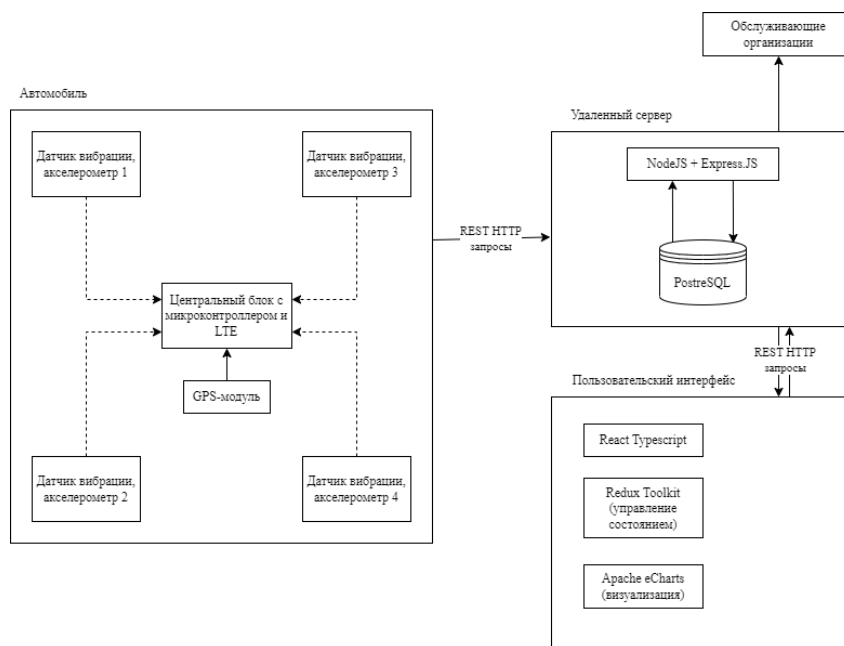


Рис. 1. Структура системы

реализовано на RESTful API. Структура системы изображена на рисунке 1.

Принцип работы системы

Разрабатываемая система работает следующим образом. При попадании машины с установленной системой в выбоину, превышении значениями вибрации с датчиков пороговых значений и продолжении движения при небольшом изменении скорости, система фиксирует повреждение и отправляет на сервер пакет данных, содержащий геопозицию, значения скорости, вибрации, акселерометра и время.

На сервере происходит анализ полученных данных, отбраковка ложных данных и автоматическое формирование запросов на ремонт в обслуживающие организации, если количество срабатываний в определенном радиусе превышает пороговое значение.

В пользовательском интерфейсе будет представлена карта с точками, где сработала система и возможностью просмотра более подробной информации о пакете данных. Также можно будет просмотреть отправленные запросы в обслуживающие организации.

Заключение

Автоматическое детектирование повреждений дорожного полотна может существенно уменьшить количество выбоин-долгожителей на дорогах, а следовательно – уменьшить аварийность для участников дорожного движения.

В работе разработана структура системы, что позволяет перейти к подбору компонентов, моделированию работы системы и разработке веб-части системы.

Список литературы

1. Jeong, Dongjun & Kim, Jua. Road Damage Detection using YOLO with Image Tiling about Multi-source Images. - 2022
2. Perwej, Dr. Yusuf & Kumar, Ankit & Kulshrestha, Neha & Srivastava, Ashish & Tripathi, Chandan. The Assay of Potholes and Road Damage Detection. International Journal of Scientific Research in Computer Science Engineering and Information Technology. - 2022. - Vol. 8. p. 202-211. 10.32628/CSEIT228135

МОДИФИКАЦИЯ ПОЛЬЗОВАТЕЛЬСКОГО ИНТЕРФЕЙСА ИЗМЕРИТЕЛЬНОГО КОМПЛЕКСА

Недошивина М.С., Язовская Д.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Изложены текущие результаты разработки интерфейса измерительного комплекса для анализа параметров различных электронных компонентов.

Введение

В лаборатории 220 МИЭМ НИУ ВШЭ организован и функционирует программно-аппаратный комплекс для получения SPICE-моделей различных электронных компонентов: биполярных и МОП-транзисторов различных видов, тиристоров (рис. 1) [1]. В комплексе используются связанные контрольно-измерительные приборы (КИП) под управлением компьютера и программное обеспечение (ПО) для идентификации параметров. К КИП, с которыми работает комплекс, относятся: источник GPD-73303S, мультиметры GEU820571 и GEU820545, осциллограф OSC AKIP-4126 и Keithley 2602. Однако, имеется ряд недостатков, касающихся интерфейса комплекса: а) он недостаточно дружелюбный, для каждого нового типа компонента и электрической характеристики требуется работа программиста для формирования дополнительных модулей, б) затруднено переключением контрольно-измерительных приборов к комплексу, в) комплекс настраивается только через текстовые файлы [2].

С целью устранения недостатков прошлого измерительного комплекса и его последующего усовершенствования в 2021 году был организован студенческий проект, который в настоящий момент продолжается. В программной среде «NI LabVIEW» [3] были разработаны новый пользовательский интерфейс и обобщенная схема измерений. В процессе разработки был существенно изменен алгоритм работы оператора с измерительным комплексом.

В процессе тестирования и анализа опыта применения разработанной новой версии комплекса был выявлен ряд логических и программных ошибок.

Анализ алгоритма работы оператора с комплексом описан в разделе I, чтобы продемонстрировать в разделе II улучшения интерфейса измерительного комплекса, позволяющие решить большинство проблем, выявленных при тестировании.



Рис. 1. Упрощенная структурная схема измерительного комплекса

1. Анализ алгоритма работы оператора комплекса

В рамках существующей системы для проведения одного измерения оператор должен пройти 6 этапов.

Сначала в настройках оператор указывает текстовый файл базы приборов, а также адреса папок для сохранения пресетов, графиков и результатов измерений. Это действие оператору необходимо выполнять каждый раз при открытии измерительного комплекса, даже если адреса папок постоянны, что замедляет работу оператора.

На следующем шаге оператору необходимо для каждого используемого канала выбрать измерительные приборы в списке приборов, информация о которых считывается с подключенных портов и отобра-

ражается на интерфейсе. Эти действия также занимают лишнее время и требуют лишних движений [4].

На третьем шаге оператор выбирает подходящий пресет – сохранённый набор настроек измерения. В пресет входят очередность обработки канала, значения источника, названия выходных переменных и режим работы для каждого канала. По своему усмотрению оператор может изменить любое из вышеперечисленных значений. Из-за обилия пресетов с похожими названиями поиск необходимого может занять время.

Наконец, после полной настройки комплекса оператор может запустить измерения вольт-амперных характеристик. После запуска в обобщенную схему измерений передаются все заданные значения. При запуске измерений не производится никакая проверка адекватности входных параметров, что может повлечь за собой некорректную работу измерительного комплекса и даже выход из строя КИП.

Последний этап работы с комплексом – это вывод графиков и сохранение результатов измерений. Для вывода графиков оператор самостоятельно выбирает переменные по оси X и по оси Y. Результаты сохраняются в файл в формате, необходимом для программы идентификации параметров SPICE-моделей [5].

Помимо указанных выше недочетов нынешнее решение интерфейса имеет некоторые другие недостатки в логике дизайна, которые выявило пользовательское тестирование: а) при повторных измерениях оператору приходится повторять весь алгоритм работы с комплексом и тратить время на задание параметров, которые не изменились с предыдущего запуска, потому что в измерительном комплексе не предусмотрено сохранение настроек; б) функциональность задания значений источника ограничена регулярным списком, что, например, не позволяет задать постоянное значение при неизменном напряжении на источнике; в) отсутствие быстрого доступа к часто используемым пресетам приводит к замедлению работы оператора. Кроме того, в интерфейсе и, как следствие, в обобщенной схеме недостаточно функциональности для измерения некоторых характеристик электронных компонентов.

2. Улучшенный интерфейс измерительного комплекса

Улучшения пользовательского интерфейса (см. рис. 2) коснулись как изменений его внешнего вида, так и логики его работы, а как следствие – логики работы оператора с интерфейсом.

а. Различные команды для работы с проектами и пресетами сведены в пользовательских меню (рис. 2 – пункт 1).

б. Для упорядочения работы оператора с измерительным комплексом введена система проектов. В проекте сохраняется важная информация о настройках процесса измерений в применении к конкретной производственной технологии: сведения об обширной базе контрольно-измерительных прибо-

ров; набор пресетов для параметров источников-измерителей, используемых в конкретной схеме измерений; пути для сохранения результатов измерений в графической и текстовой форме.

в. Из большого числа пресетов, относящихся к различным типам электронных компонентов, измеряемых характеристик, производственных технологий, оператор может выбрать короткий список избранных пресетов (рис. 2 – пункт 2), относящихся к текущему сеансу измерений. Список избранных пресетов сохраняется в проекте.

г. При повторяющихся измерениях множества интегральных транзисторов, расположенных на одном полупроводниковом кристалле, удалось значительно сократить время, которое тратилось на перезаполнение свойств транзистора (порядкового номера, геометрических размеров: ширины и длины, полярности) при каждой смене объекта измерений, т.е. каждые несколько минут, – за счёт возможности подключения предварительно заполняемого каталога свойств транзисторов (рис. 2 – пункт 3).

д. Введена возможность задавать значения источников сигнала в различной форме: постоянное значение, нерегулярный список с перечислением значений (рис. 2 – пункт 4).

е. Если необходимо подавать управляющие электрические сигналы на какие-либо выводы измеряемого компонента, а измерять выходные сигналы на них не нужно, то в этом случае можно использовать сокращённые версии измерительных каналов, содержащие только источники питания (рис. 2 – пункт 5). Это приводит к некоторым изменениям в обобщенной схеме измерений.

ж. У всех каналов, кроме двух основных, введена возможность их отключения (рис. 2 – пункт 6) при необходимости.

и. Для улучшения наглядности интерфейса надписи на кнопках были заменены на значки; кроме того, введены всплывающие подсказки с дополнительной информацией при наведении на них (рис. 2 – пункт 7). Это позволяет снизить визуальную нагрузку в интерфейсе, ускоряет работу оператора с ним, улучшает понимание возможностей интерфейса [6].

к. Усовершенствована система безопасности проведения измерений (рис. 2 – пункт 8). Во все модули комплекса, где это необходимо (например, при инициализации приборов, измерениях и деинициализации приборов), внедрена система обнаружения, фиксации ошибок при взаимодействии с КИП. Предложенный механизм дает возможность оповещать оператора о возникновении ошибок и позволяет ему оперативно реагировать на них. Это снижает вероятность выведения аппаратуры и измеряемых компонентов из строя,

л. Разработана система проверки на адекватность данных, вводимых оператором вручную в поля пользовательского интерфейса: очередность использования измерительных каналов, значений сигналов источников, температуры и значений измеряемых сигналов. Рекомендательная система указывает оператору на причины проблемы.

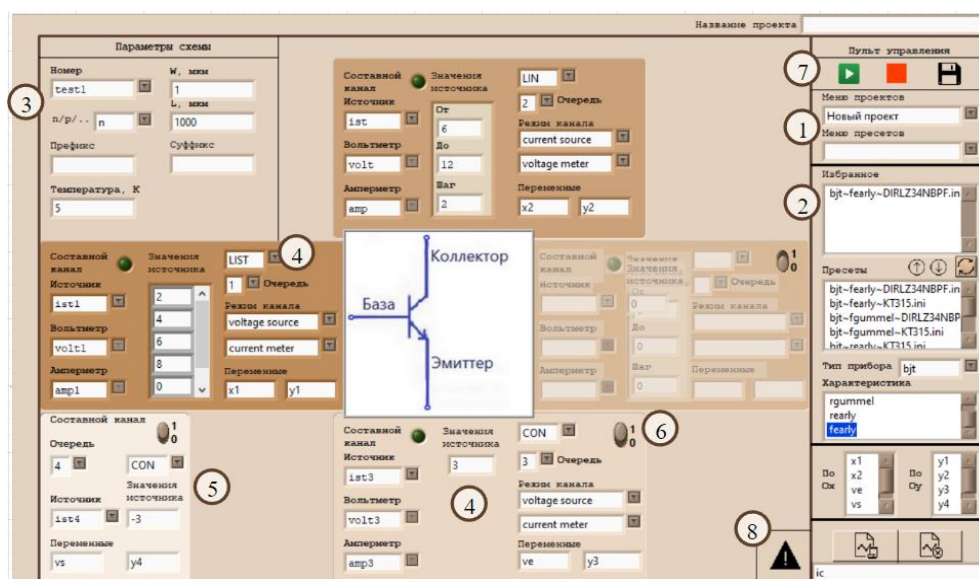


Рис. 2. Улучшенный пользовательский интерфейс; пронумерованные элементы интерфейса описаны в сопровождающем тексте

Заключение

По результатам обширного тестирования измерительного комплекса, разработанного на предыдущем этапе выполнения проекта, был выявлен ряд недостатков, затрудняющих планомерное выполнение работ по измерению электрических характеристик партий электронных компонентов.

Для исправления сложившейся ситуации были реализованы некоторые улучшения пользовательского интерфейса, касающиеся как изменений его внешнего вида, так и логики его работы.

В результате существенно возросли возможности оператора по управлению процессом измерений, снижена вероятность влияния человеческого фактора, увеличена скорость выполнения работы.

Развитие проекта предусматривает внедрение системы активных подсказок оператору при использовании интерфейса. Разработанная обобщенная схема измерений будет распространена на более широкий круг различных характеристик и типов электронных компонентов.

Список литературы

1. K. O. Petrosyants, L. M. Sambursky, I. A. Khartanov and M. R. Ismail-zade, «Generalized test automation method for MOSFET's including characteristics measurements and model parameters extraction for aerospace applications», 2017 IEEE East-West Design & Test Symposium (EWDTS), Novi Sad, Serbia, 2017, pp. 1-8, doi: 10.1109/EWDTS.2017.8110055.
2. Копытин И.А., Недошивина М.С., Язовская Д.А., «Разработка универсального интерфейса измерительного комплекса», Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В. Арменского, 2022, с. 233-235. [Электронный ресурс]. URL:

<https://miem.hse.ru/data/2022/09/02/1700621959/MIEM-HSE-2022-final.pdf>. (дата обращения: 17.01.2023)

3. LabVIEW Support [Электронный ресурс]. URL:

<https://www.ni.com/ru-ru/support/downloads/software-products/download.labview.html#443274> (дата обращения: 21.01.2023)

4. Вумек Дж., Джонс Д. «Бережливое производство: Как избавиться от потерь и добиться процветания вашей компании». – Пер. с англ. – 7-е изд. – М.: Альпина Паблишер, 2013 – 472 с.

5. IC-CAP 2006, User's Guide, Agilent 85190A

6. Купер А., Рейман Р., Кронин Д. «Алан Купер об интерфейсе. Основы проектирования взаимодействия». – Пер. с англ. – СПб.: Символ-Плюс, 2009 – 688 с.

СРАВНЕНИЕ СПОСОБОВ РАЗРАБОТКИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ САМООБСЛУЖИВАНИЯ ДЛЯ ИНВЕНТАРИЗАЦИИ ДЛЯ БИБЛИОТЕКИ НА ОСНОВЕ RFID-МЕТОК И QR-КОДОВ

Чуйко Д.О., Кретова А.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В ходе работы исследованы методы создания автоматизированных систем самообслуживания для проведения инвентаризации, приведены некоторые существующие модели, а также описан прототип разрабатываемой системы, необходимые для создания технические компоненты и программное обеспе-

чение. В заключении работы приведено сравнение разрабатываемой системы с существующими, а также возможные зоны роста.

Введение

В современном мире многие библиотеки перешли в электронный формат, и появилось множество сайтов с отзывами, позволяющими легко выбрать книгу, которую будете читать следующей. Таким образом, библиотекари сейчас по большей части выполняют роль «принеси и запиши, что книга была взята или сдана». Из этого можно сделать вывод, что вместо библиотекарей стоило бы сделать автоматизированную систему самообслуживания для посетителей библиотеки. Из-за ряда проблем в традиционной работе библиотек её необходимо улучшить и автоматизировать. Таким образом, выбранная тема разработки является актуальной и охватывает множество бюджетных организаций, работающих децентрализованно и по устаревшей методике. Что касается практической значимости, то она заключается в том, что разработанная система позволит автоматизировать и ускорить работу любой библиотеки, а также обеспечить сохранность имущества.

В данный момент подобные системы уже существуют и реализованы в библиотеке Высшей Школы Экономики на ул. Мясницкой или на Покровском бульваре, но они являются довольно дорогостоящими и сложными в установке. Поэтому подобные системы не подходят для обычных библиотек, не обладающих большим бюджетом. Кроме того, на постоянной основе необходима сторонняя организация, обеспечивающая поддержку работоспособности такой системы.

Предлагаемое решение позволит в короткий срок установить систему самообслуживания, внести все предметы, подлежащие инвентаризации, в базу данных и не потребует для этого большого количества вычислительных ресурсов и дорогостоящего оборудования. Кроме того, будет возможность создать механизм защиты вещей от краж, установив на выходе из помещения рамки, которые будут сигнализировать при несанкционированном выносе каких-либо вещей.

Нельзя не сказать о необходимости инвентаризации для того, чтобы постоянно отслеживать количество находящихся в распоряжении книг, анализировать интересы читателей и своевременно покупать необходимые книги. Инвентаризация представляет собой достаточно трудоемкий процесс, который связан с большими объемами и разнообразной номенклатурой материальных ценностей, а также недостатками в обеспечении контроля за их сохранностью [1].

Основным источником для разработки системы служат существующие научные статьи, доклады с научных конференций, а также документации на используемые в ходе работы программные компоненты. В одной из статей была продумана подобная система для инвентаризации с помощью штрих-кодов. [2]. Архитектура системы и взаимодействие с посе-

чителями из статьи послужило основой разрабатываемой системы. К тому же довольно объемной и полезной является другая научная статья, где описана технология применения RFID-меток для улучшения работы библиотеки [3]. Благодаря данной работе был продуман функционал разрабатываемой системы, в частности функции взятия книг с помощью считывателя RFID-меток, а также поиск книг на полках с помощью специального оборудования. Важность инвентаризации, являющейся основным средством контроля при приемке-передаче материальных ценностей, описана в другой работе, которая к тому же позволяет погрузиться в предметную область. [4] Разработка системы контроля на основе интегрированного программно-аппаратного комплекса (ИПАК) Arduino описана в статье научно-технического журнала и предоставляет информацию о возможностях и ограничениях ИПАК и особенностях взаимодействия с ним. [5] Об опыте использования RFID-меток, возможных трудностях и методах их решения можно узнать из другой научной статьи. [6] На основании этих статей при разработке собственной системы будут учтены факторы внешней окружающей среды и помех, которые могут влиять на правильность считывания меток, будет отдан приоритет при выборе оборудования в пользу того, которое уже применялось ранее для реализации схожих систем, а также получена информация об опыте использования подобных автоматизированных систем самообслуживания и учтены все допущенные ошибки, такие как неправильная настройка сканеров и считывателей RFID-меток, неправильная идентификация книг с помощью QR-кодов и меток, а также неправильная архитектура системы, из-за которой могли появляться дублирующиеся книги в базе данных.

Прототип разрабатываемой автоматизированной системы самообслуживания, необходимые компоненты и программное обеспечение

Для разработки и реализации прототипа потребуются технические компоненты и программное обеспечение. Чтобы снизить стоимость разрабатываемого решения будет использовано по большей части бесплатное программное обеспечение, а из комплектующих – наиболее популярные и выгодные по соотношению цены и качества.

Планируется параллельно реализовать две системы для инвентаризации. Первая не будет включать в себя систему защиты от несанкционированного выноса предметов из помещения. Она основана на использовании QR-кодов, которые и позволяют автоматизированно вести инвентаризацию. Вторая система будет основана на RFID-наклейках, которые в случае выноса из помещения будут пищать при прохождении антенн в том случае, если не было получено разрешение на вынос этих предметов. Пример расположения такой метки представлен на рис. 1.

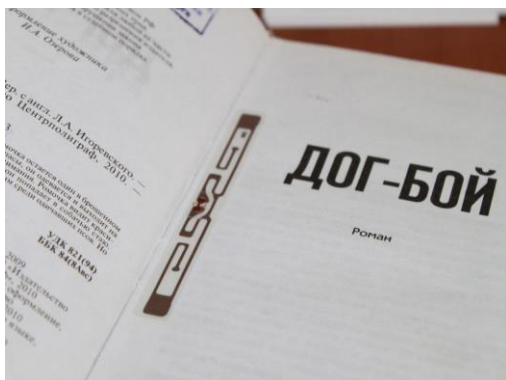


Рис. 1. RFID-метка в книге

Теперь более подробно рассмотрим каждую из этих систем, а также необходимое программное обеспечение и оборудование. В качестве места для реализации выберем библиотеку с хранящейся в ней книгами. Система будет реализована на веб-сервере.

Перейдём к описанию первой системы. Сперва книга поступает в библиотеку. Необходимо добавить её в базу данных. На сайте необходимо авторизоваться в качестве администратора, перейти на страницу «Добавить книгу». Далее необходимо ввести всю информацию о книге, включая её автора, название, жанр и описание. Сайт генерирует QR-код, который необходимо распечатать на клеевой бумаге и приклеить к книге. В это же время книга автоматически заносится в базу данных вместе с другими книгами. Важным условием для эффективной работы является наличие идентификационных номеров на всех проверяемых документах [7].

Таким образом, при взятии книги можно будет понять, что взята именно эта книга, а не какая-то с аналогичным названием и автором. После регистрации книги в базе данных её можно ставить на полку. Опционально можно разметить все полки в помещении и ставить на конкретное место, а затем сохранять его в базе данных, чтобы при необходимости можно было легко найти предмет. Но у подобного решения может быть и минус – студенты, желающие почитать книгу, могут взять её с одного места, а поставить на другое. В таком случае система хранения является неудобной и будет только сбивать с толку. Если студент решает взять книгу с собой, то ему необходимо подойти к терминалу самообслуживания, которым может выступать обычный компьютер с веб-камерой или сканером QR-кодов, отсканировать код, как на рис. 2, а дальше идентифицировать себя в системе – авторизовавшись в личном кабинете библиотеки. Процесс регистрации может свестись к минимуму – фотография студенческого билета, электронная почта и ФИО. Студенту будет предоставлен уникальный идентификационный номер, по которому он будет брать книги.

При взятии книги помимо выдачи указывается срок, в который необходимо вернуть книгу (по умолчанию срок будет равен 2 неделям). Продлить его можно будет в личном кабинете самим студентом. Если в указанный срок книга не будет возвра-

щена, то на почту администратору приходит сообщение с информацией о студенте и его электронной почтой, на которую можно будет отправить сообщение с напоминанием о необходимости вернуть книгу. Система также автоматически может напоминать о необходимости возвращении книги. Это может быть реализовано с помощью бота в социальной сети Telegram или сообщениями на корпоративную почту.



Рис. 2. Сканирование QR-кода и вывод номера книги

Преимуществом такой системы является то, что процесс выдачи книги не требует нахождения рядом библиотекаря, а весь процесс сопровождается инструкциями. Кроме того, выдача является довольно быстрой и не займет больше минуты, если не нужно будет регистрироваться при первом посещении или студент не забудет пароль.

Кроме того, можно создать систему идентификации с помощью RFID-меток – это и является основой второй разрабатываемой системы. В основе популярности и стремительного развития технологии RFID лежит целый ряд факторов. Во-первых, это надежность. Радиочастотные метки долговечнее продуктов, на которых они размещены, существуют различные технологии обеспечения их термостойкости, водостойкости и ударопрочности. Во-вторых, технология RFID не требует прямой видимости метки и позволяет считывать информацию одновременно с большого количества таких меток [8]. Процесс инвентаризации очень похож на первый за исключением того, что основой служат не QR-коды, а RFID-метки. При поступлении книги на неё клеится уникальная RFID-метка, которая далее идентифицирует книгу в системе. Технология максимально упрощает процесс взятия книги, ведь всё, что нужно сделать – это поднести книгу с заранее наклеенной на неё меткой к считывателю RFID-меток. Книга появится на экране с информацией о ней: автором, названием, описанием и статусом бронирования. Если книга доступна для бронирования, то необходимо выбрать соответствующую опцию и приложить свой пропуск в здание университета, который также основан на технологии RFID-меток и позволит уникально идентифицировать студента.

Вторая система является более удобной и простой в использовании, нежели первая, но при интеграции в существующую корпоративную систему могут возникнуть проблемы с политикой конфиденциальности, а также передачей информации о пропусках и

их владельцев. К тому же вторая система позволяет изменять статус бронирования и в том случае, если книга взята в аренду, студент сможет спокойно покинуть помещения, пройдя через рамки-антенны. В случае, когда книга была взята несанкционированно и будет находиться в зоне действия рамок, раздастся звуковой сигнал, означающий то, что книга без доступа «на вынос» покидает библиотеку.

Программное обеспечение, необходимое для реализации подобных систем включает в себя выделенный виртуальный сервер или хостинг, а также среду разработки. При необходимости могут быть использованы фреймворки. Техническое оборудование включает в себя RFID-метки, считыватель меток, антенны, клейкую бумагу, сканер QR-кодов или веб-камеру, а также персональный компьютер. Примером считывателя RFID-меток может служить микроконтроллер Arduino UNO с RFID-модулем чтения/записи RF522, изображенный на рис. 3.

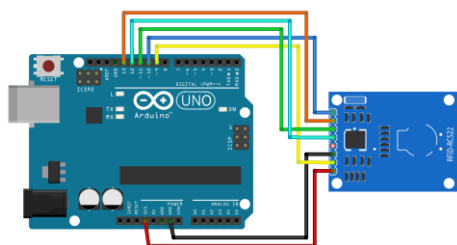


Рис.3. Микроконтроллер Arduino UNO с RFID-модулем RF522

Сравнение разрабатываемой системы с существующими, зоны роста

На данный момент уже существует система автоматизированной выдачи книг, которая организована следующим образом. На каждой книжке есть штрих-код, который позволяет определить, какая книга взята. К сожалению, он не дает возможность точно идентифицировать книгу, а лишь получить её название и автора. Далее посетитель библиотеки приходит с этой книгой к библиотекарю, который спрашивает фамилию, имя и отчество и закрепляет за этим читателем эту книгу, отсканировав штрих-код на ней с помощью сканера. Такая система тоже является более удобной по сравнению с прежней, где всю информацию приходилось записывать вручную, тратить время на поиск личной книжки посетителя, а также записывать полностью название и автора взятой книги. Подобное решение экономит время обслуживания и упрощает процедуру получения книги, но все равно не полностью автоматизирует её. Из-за этого могут случиться ошибки при записи названия книги, указании личных данных читателя, а также нет возможности отследить оставшиеся в библиотеке книги, их количество, состояние и местонахождение.

Другой реализованной системой является автоматизированная библиотека в Томском Государственном университете. Там процесс является более автоматизированным, поскольку уже включает использование RFID-меток, которые позволяют точно иден-

тифицировать книгу. При прикладывании книги выводится вся информация о ней. Но пока не было принято решение отказаться от библиотекарей, поскольку выдача книг без них является рискованной в силу отсутствия надежной системы безопасности, позволяющей моментально заметить и сообщить о краже. При установке рамок на входе и выходе можно было бы отслеживать, не проносится ли книга мимо библиотекаря из помещения библиотеки. Подобная задумка уже есть и находится в стадии разработки, но точных сроков по её реализации не указано. Кроме того, данная система является довольно дорогостоящей, а на периодическое обслуживание приходится тратить не малую сумму денег в связи с чем возникает вопрос, стоит ли в целом реализовывать подобные системы, если на поддержку системы тратится такая сумма, на которую можно было бы купить пару десятков хороших книг и увеличить библиотечный фонд.

Несмотря на то, что некоторые подобные системы уже разработаны и функционируют, они все ещё являются дорогостоящими и сложны в установке и запуске. Тем не менее, нельзя не оценить важность таких автоматизированных систем по сравнению с теми, где для ведения учёта необходим человек.

Использование автоматизированных систем с терминалами самообслуживания позволяет:

- автоматизировать процессы выдачи и возврата;
- защитить фонд от несанкционированного выноса;
- упростить поиск;
- проводить инвентаризацию с минимальными затратами временных и человеческих ресурсов;
- расширить часы работы библиотеки;
- повысить качество обслуживания;
- исключить возможность человеческой ошибки при работе с книжным фондом или другим имуществом [9].

Одним из хороших решений для улучшения разрабатываемой системы является создание датчиков позиционирования предметов в пространстве. На территории библиотеки необходимо размещение приемочной базы, которая будет принимать сигнал с каждой метки и посредством кабельного подключения к сети Интернет, передавать данные на сервер для обработки и вычисления позиционирования книги в пространстве. Вся входящая информация на сервер сохраняется и может передаваться пользователям системы [10]. Такое улучшение упростит и снизит временные затраты на поиск книг.

Помимо процесса хранения и выдачи книг стоит продумать систему инвентаризации, которая периодически производится. Так, при выборе опции инвентаризации можно будет сканировать все книги поочередно, а в конце посмотреть, какие книги не были учтены – возможно, они где-то затерялись или были пропущены. Это экономит время и силы на проведение такого масштабного и важного процесса.

Возможно введение рекомендательного сервиса, когда по возвращению книги будут предложены книги того же автора или того же жанра. Такая система упростит поиск следующей книги для чтения. А по-

сколько книги будут предлагаться из имеющихся в наличии, то не нужно будет думать по поводу её доступности.

Заключение

В ходе данной работы была рассмотрена область автоматизации системы выдачи и приёма книг в библиотеках. Были приведены и описаны два способа реализации и работы автоматизированной системы самообслуживания для библиотеки. Один из них заключается в реализации системы с помощью RFID-технологий, а другой основан на QR-кодах. Приведены преимущества и недостатки подобных способов реализации, а также возможности, которые открываются при использовании автоматизированной системы с терминалами самообслуживания. Подтвердить актуальность и практическую значимость работы можно следующими фактами. Во-первых, подобная система значительно упрощает процесс получения и возврата книг для читателей библиотеки, что сокращает время ожидания очереди и делает процесс более удобным для посетителей. Во-вторых, автоматизация процессов повысит эффективность работы библиотекарей, освободив их от рутинных операций, что позволит сконцентрироваться на важных задачах. В третьих, автоматизация библиотечных процессов может помочь библиотекам сократить затраты на персонал и обслуживания, что в свою очередь позволит увеличить бюджет на покупку новой литературы или улучшить доступность, включая людей с ограниченными возможностями.

Список литературы

1. Мякина Л.Н. Инвентаризация товаров // Бухгалтер и закон. Общество с ограниченной ответственностью «Издательский дом ФИНАНСЫ и КРЕДИТ», 2005. № 11 (93).
2. Аветисов М.А., Евгеньева Б.С., Косикова Н.В. Инвентаризации объектов библиотечного фонда ЦНСХБ с помощью технологий штриховых кодов // Культура: теория и практика. 2016. №3 (12).
3. Морев Владимир Алексеевич, Тимошук Максим Олегович Применение радиочастотных систем (RFID) в библиотечном деле (на примере научной библиотеки Национального Исследовательского Томского Государственного Университета) // Научный результат. Технологии бизнеса и сервиса. 2020. №3.
4. Л.Н. Мякина Инвентаризация товаров // Бухгалтер и закон. 2005. №11 (93).
5. Шарманов В.В., Симанкина Т.Л., Горбачев И.А. Идентификация местоположения работника на объекте строительства с помощью цифрового двойника // Вестник Научного центра. 2022. №3.
6. Черепков Сергей Технология RFID - радиочастотная идентификация. Опыт использования и перспективные направления // Компоненты и Технологии. 2005. №53.
7. Инвентаризации объектов библиотечного фонда ЦНСХБ с помощью технологий штриховых кодов [Electronic resource]. URL:

<https://www.elibrary.ru/item.asp?id=27207428> (accessed: 13.12.2022).

8. Сергей Ч. Технология RFID - радиочастотная идентификация. Опыт использования и перспективные направления // Компоненты и Технологии. Общество с ограниченной ответственностью «Издательство Файнстрит», 2005. № 53.

9. Идентификация местоположения работника на объекте строительства с помощью цифрового двойника [Electronic resource]. URL: <https://www.elibrary.ru/item.asp?id=49487910> (accessed: 13.12.2022).

10. Алексеевич М.В., Олегович Т.М. Применение радиочастотных систем (RFID) в библиотечном деле (на примере научной библиотеки Томского государственного университета) // Научный результат. Технологии бизнеса и сервиса, 2020. Vol. 6, № 3.

КОНЦЕПЦИЯ СИСТЕМЫ ОБНАРУЖЕНИЯ ИСТОЧНИКОВ РАДИОСИГНАЛА НА БАЗЕ БПЛА

Макаренко Д.В., Матвеев Д.С.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Предложена концепция системы обнаружения сигнала несанкционированного управления беспилотных летательных аппаратов с воздуха. Система включает в себя беспилотный летательный аппарат с установленными приёмо-передающим и сканирующим оборудованием, а также аппаратурой дистанционного управления. Разрабатываемую систему предполагается использовать для выявления факта несанкционированного управления беспилотных летательных аппаратов. Приведена структура аппаратной части системы и алгоритм поиска сигнала.

Введение

Стремительное развитие технологий беспилотных летательных аппаратов (БПЛА), привело к тому, что БПЛА стали применяться для противоправной деятельности, такой как террористические атаки, направленные в сторону людей, объектов инфраструктуры, частной собственности [1-2]. Также, способность БПЛА нести полезную нагрузку может быть использована, например, для организации оборота запрещенных веществ. В Российской Федерации с 2016 года некоторые типы беспилотных воздушных судов (БВС) подлежат обязательной сертификации. На момент написания статьи БВС весом от 150 г. должны проходить государственный учёт и получать разрешения на полёты [3].

Актуальность разработки подтверждают введённые указы о запрете использования БВС на территории ряда субъектов РФ [4]. Таким образом задача

локации операторов нелегальных БВС становится ещё более востребованной. Целью работы является обеспечение возможности поиска несанкционированного управления БПЛА. Для такого устройства особое значение приобретают такие показатели системы как скорость развёртывания системы, габариты, масса, способность совершать поиск независимо от времени суток и года, в условиях пересечённой местности, в условиях городской застройки, возможность визуально оценивать предполагаемый объект излучения.

В данной работе будет описана концепция системы и алгоритм поиска направления источника радиосигнала.

Структура аппаратной части реализуемого решения

Система состоит из следующих компонентов: квадрокоптер, направленная антенна, поворотный механизм с камерой, система Head tracker, одноплатный компьютер Orange Pi, Software-defined radio (SDR), шлем и аппаратура для управления квадрокоптером.

Существующие станции для определения местоположения источника сигнала в основном используют угломерные системы. Известно, что с помощью одного неподвижного пеленгатора можно определить только азимут направления на источник радиозлучения. Для определения местоположения источника радиозлучения необходимо иметь, по крайней мере, два радиопеленгатора, удаленные друг от друга на достаточное расстояние [5]. Однако, отличительной особенностью данной работы является использование подвижной сканирующей системы для определения направления источника сигнала, на первоначальном этапе разработки, и обеспечение возможности определения его точных координат в дальнейшем.

Беспилотный аппарат как правило имеет сигнал управления на частоте сигнала 2,4 ГГц. Для детектирования частоты предлагается использовать спиральный тип антенн (Helix). Такая антенна имеет циркулярную поляризацию с диаграммой направленности 30 градусов и усилением сигнала около 12,5 dB. Конструкция антенны имеет относительно малые габариты и вес, что подходит для установки её на раму квадрокоптера. Также антенна имеет изоляцию в виде дефлектора для обеспечения достоверности получаемых в ходе детектирования показаний. В дальнейшем, для расширения полосы частот возможно заменить антенну. Для обеспечения подвижности антенны планируется разработка модели крепления с сервоприводом и системой управления вращением Head tracker.

Одноплатный компьютер и SDR-приемник, которые обрабатывают получаемые данные, размещаются в закрытом корпусе, изготавливаемом с использованием FDM-технологии 3D печати.

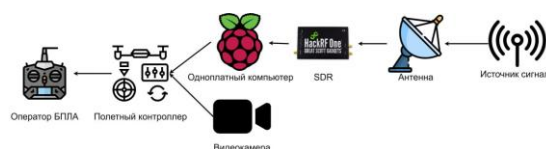


Рис. 1. Архитектура системы

Бортовая сеть на БПЛА запитывается от аккумулятора 25В. Далее через преобразователь напряжение понижается до 5В, от которых питается одноплатный компьютер и SDR-приёмник.

Алгоритм поиска

Для осуществления поиска радиосигнала используется программно-определяемое радио HackRF One и утилита `hackrf_sweep` [7], позволяющая получать данные о мощности сигналов на заданном диапазоне с частотой сканирования 8ГГц/с и выводить результаты измерений в виде значения, разделенного запятыми (csv формат). Далее, данные через usb-порт передаются в компактный одноплатный компьютер, на котором установлена unix-подобная операционная система с запущенной контейнеризацией Docker [8]. Внутри контейнера выполняется программа на языке программирования Python, позволяющая обрабатывать информацию о сигналах, и реляционная база данных, для сохранения обработанных значений. Собранные данные агрегируются с информацией, полученной с GPS-модуля, барометра и цифрового компаса, установленных на БПЛА. Полученный набор данных является необходимым и достаточным для определения вектора направления на источник радиосигнала.

Для большего удобства управления БПЛА операторы используют технологию OSD [9] (On-screen display). Она позволяет наложить текстовые данные поверх видеопотока, захватываемого камерой. Наложение происходит в полетном контроллере беспилотника. В рамках предложенной концепции к стандартным метрикам базовой конфигурации полетного контроллера добавлены новые значения, получаемые с программно-определяемого радио и обработанные одноплатным компьютером.

Заключение

Таким образом, в данной работе описаны основные функциональные возможности системы, представлен состав аппаратной части и показана схема их взаимодействия. Предложенный алгоритм в комплексе с используемым приемо-передающим и сканирующим оборудованием позволит выявлять направление источника внешнего радиосигнала.

В дальнейшем планируется доработка системы для обеспечения возможности определения точных координат источника внешнего радиосигнала.

Список литературы

1. Мировые инциденты и происшествия с дронами и БПЛА [Электронный ресурс]. URL: <https://www.rub-in.ru/intsidenty-s>

dronami/?ysclid=ldcjrduq999117556 (дата обращения: 20.01.2023).

2. Каляев, А. И. Комплекс обнаружения и поражения БПЛА-нарушителей с помощью БПЛА-перехватчиков / А. И. Каляев, Я. С. Коровин // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. – 2021. – № 3-4(153-154). – С. 101-107. – EDN GMUUD.

3. "Воздушный кодекс Российской Федерации" от 19.03.1997 N 60-ФЗ (ред. от 19.12.2022).

4. Полный список запретов по регионам на полёты дронов/квадрокоптеров/БПЛА в России в 2022/2023 [Электронный ресурс]. URL: <https://coptertime.ru/reviews/articles/no-fly-2022/?ysclid=ldankm7p1f561811133#31> (дата обращения: 23.01.2023).

5. Портативная система радиомониторинга и определения местоположения источников радиоизлучения [Электронный ресурс]. URL: <http://www.bnti.ru/showart.asp?aid=668&lvl=04.01.01.01.&ysclid=lc5uk72the217927419> (дата обращения: 25.12.2022).

6. Hackrf_sweep [Электронный ресурс]. URL: https://hackrf.readthedocs.io/en/latest/hackrf_sweep.html (дата обращения: 19.01.2023).

7. Docker: Accelerated, Containerized Application Development [Электронный ресурс]. URL: <https://www.docker.com/> (дата обращения: 19.01.2023).

8. OSD– RSearch [Электронный ресурс]. URL: <https://rsearch.ru/wiki/OSD> (дата обращения: 15.01.2023).

СИСТЕМА ОТСЛЕЖИВАНИЯ ЧЕЛОВЕКОПОТОКА В РЕАЛЬНОМ ВРЕМЕНИ

*Захваткина В.Ф.¹, Дубовикова С.А.²,
Некрасов Ф.И.¹*

*Национальный исследовательский университет
«Высшая школа экономики»,*

¹департамент электронной инженерии,

*²департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе рассматривается реализация системы отслеживания человекопотока в реальном времени. Определяется целесообразность пересылки необработанных данных с помощью IPv6 over BLE. Предлагается алгоритм обработки данных на микроконтроллере для определения количества и направления проходящих через дверной проём людей.

Ключевые слова: IPv6 over BLE, ИК-сенсоры, человекопоток.

Введение

Изучение трафика и конверсии точек розничной торговли, оптимизация маршрутов и интервала движения общественного транспорта, улучшение работы подсистем умных городов – лишь небольшая

часть того, где может быть использовано отслеживание человекопотока.

Существующие на рынке системы отслеживания человекопотока дорогие и недостаточно точные. Для повсеместного внедрения подобных решений необходимо разработать архитектуру, позволяющую снизить стоимость внедрения и обслуживания, сохранив при этом достаточную точность и энергоэффективность.

Целью данной работы является улучшения процесса обработки и передачи данных системы отслеживания человекопотока с использованием инфракрасных сенсоров-матриц.

Новизна заключается в возможности определения количества людей, проходящих под датчиком, на основе температурной картины, полученной с ИК-датчика низкого разрешения.

В первой части работы проведена оценка пропускной способности и энергоэффективности IPv6 over BLE для определения целесообразности обработки данных на оконечных устройствах. Далее описывается алгоритм, позволяющий фиксировать направление движения одного человека. Наконец, в третьей части приводится доработка алгоритма обработки данных, позволяющая определять не только направление проходящих людей, но и с высокой точностью устанавливать их количество.

Основными методами проведения исследования являются изучение и анализ научно-технической литературы и разработка.

IPv6 over BLE

Bluetooth Low Energy (BLE) на сегодняшний день является наиболее распространенной беспроводной технологией с низким энергопотреблением.

Низкое энергопотребление передачи по BLE значительно повышает время автономной работы устройств, что положительно влияет на ее популярность для IoT. Также привлекательность BLE для реализации IoT решений обуславливается, помимо энергоэффективности, появлением стандартизации IPv6 по Bluetooth Low Energy (6LoBLE), которая определяет использование BLE в качестве канального уровня в сетях 6LoWPAN на основе IPv6. Это позволяет осуществлять крупномасштабные развертывания с несколькими переходами, обеспечивая высокую производительность BLE, а также совместимость и открытость сети [1].

Кроме того, стоимость устройств и модулей, поддерживающих BLE довольно мала.

Однако стоит отметить, что энергоэффективность BLE во многом обуславливается тем, что технология в первую очередь предназначена для нечастой пересылки небольшого количества данных [2].

Потребляемая мощность варьируется в пределах от 0,01 до 0,50 Вт (в зависимости от варианта использования). В случае построения ячеистой топологии с пересылками через узлы-координаторы будет проходить большое количество данных, что заметно скажется на энергоэффективности сети, а значит и снизит время автономной работы системы.

Рассмотрим количество данных, отправляемых в случае обработки вне микроконтроллера:

1. Для записи значений температуры используем int8_t , размер которого составляет 1 байт (диапазон достаточен для передачи значения температуры). Значит для описания одного паттерна потребуется 64 байта;

2. Допустим, что модуль детектирования осуществляет отправку 1 раз в минуту;

3. Для описания одного факта входа/выхода используется 6 кадров или 384 байта (3072 бита).

С множества оконечных устройств данные пересылаются координатору сети, что приводит к сокращению количества узлов, которые можно объединить в одну сеть без снижения времени автономной работы системы, так как для отправки всех данных, полученных с датчиков даже низкого разрешения, требуется значительная пропускная способность и мощность. Исходя из этого, целесообразнее обрабатывать данные непосредственно на микроконтроллере и передавать по сети только результат.

Далее предлагается алгоритм обработки получаемых с ИК-датчиков температурных паттернов.

Распознавание одного человека

Базовым требованием к системам отслеживания человекопотока является способность фиксирования факта входа и выхода из контролируемой зоны. Для этого предлагается использовать алгоритм, описанный в работе [3].

После получения данных о температурной картине с ИК-датчика разрешением 8×8 определяются \min и \max значения, которые необходимы для нормализации изображения. Нормализация осуществляется по формуле (1):

$$I_t(x, y) = \frac{F_t(x, y) - \min(F_t)}{\max(F_t) - \min(F_t)}, \quad (1)$$

где $I_t(x, y)$ и $F_t(x, y)$ – нормализованное и исходное значения пикселя с координатами (x, y) , $\min(F_t)$ и $\max(F_t)$ – минимальное и максимальное значения среди всех пикселей текущего кадра.

Далее изображение условно разделяется на 4 столбца размером 8×2 (2):

$$I_{t,i} = I_t(x, y) \text{ for } (i-1) * 2 + 1 \leq x \leq (i-1) * 2 + 2, 1 \leq y \leq 8, i \in [1, 4], \quad (2)$$

где $I_{t,i}$ – нормализованное значение пикселя, принадлежащего столбцу с номером i .

Для каждого из столбцов вычисляется мощность по формуле (3):

$$P_{(t,i)} = \frac{\sum I_{t,i}^2}{16}, \quad (3)$$

где $P_{(t,i)}$ – искомая мощность столбца под номером i , $I_{t,i}$ – нормализованное значение пикселя, принадлежащего столбцу с номером i .

Затем по данным мощностей с пяти последних кадров по формуле (4) определяется пороговое значение T_t , при котором столбец считается активным, т.е. в данный момент времени кто-то идет под датчиком.

$$T_t = k \sum_{i=1}^4 \sum_{t'=t-4}^t P_{t',i}, \quad (4)$$

где $P_{t',i}$ – мощности всех столбцов с последних пяти кадров, k – некоторый нормализующий коэффициент (подбирается экспериментально).

В текущем кадре находится столбец с максимальным значением мощности, которое должно быть $> T$. Столбец считается активным, и его номер i сохраняется. Если такого нет, то номеру активного столбца присваивается значение $i = -1$.

Теперь необходимо определить направление движения человека (вход или выход из помещения).

Для этого сохраняются номера первого (*first*) и последнего (*last*) задействованных столбцов и размер всей последовательности номеров активных столбцов (*size*).

При входе человека в поле датчика будут активны крайние столбцы с номерами “0” или “3”. Если вхождение первое, т.е. человек только что попал в поле видимости, то номер активного столбца запоминается как *first* и размер последовательности увеличивается.

Если человек уже пребывает некоторое время в поле датчика и все еще в нём находится, т.е. какой-то из столбцов активен, номер этого столбца записывается как *last* и размер последовательности увеличивается.

При начале выхода человека из кадра вновь будут активны крайние столбцы с номерами “0” или “3”, а при полном выходе номер примет значение “-1”. В таком случае определяется направление движения прошедшего человека, исходя из сравнения значений *first*, *last* и *size*, и на основании Таблицы 1 принимается решение о направлении.

Таблица 1. Соотношение полученных значений с направлением

first	last	size	Направление
3	0	4	I
0	3	4	O
3	3	> 4	ИНО
0	0	> 4	ОНИ
3	0	> 4	ИНИ
0	3	> 4	ОНО

Значения направлений расшифровываются следующим образом:

- “Т” и “О” – вход с одной стороны датчика и выход с противоположной (вход в помещение или выход из него);

- “ИНО” и “ОНИ” – вход с одной стороны датчика и выход с неё же (человек начинает выходить/выходит, но не заканчивает действие, а развивается обратно);

- “ИН” и “ОНО” – вход с одной стороны датчика и выход с противоположной (в случае остановки или медленного перемещения человека в поле датчика).

В некоторых случаях определения одного конкретного человека может быть недостаточно, например при одновременном или плотном прохождении внутрь контролируемой зоны нескольких людей.

Распознавание нескольких людей

На основе работы [4] была проведена доработка представленного выше алгоритма.

Для улучшения точности алгоритма за счет определения не только направления, но и количества проходящих людей предлагается следующий порядок действий:

В каждом последовательном кадре находится максимальное значение температуры. Во всех направлениях от пикселя с максимальным значением запоминаются координаты соседей, отстоящих от максимального на d (подбирается экспериментально). Таким образом формируется «пятно».

Полученное изображение условно делится на столбцы по формуле (2) и для получившегося пятна определяется мощность по формуле (3). По значению мощности пятна в каждом из столбцов определяется к какому из них оно относится.

Процесс повторяется для оставшихся пикселей N (подбирается экспериментально) число раз. Для тепловизора с разрешением 8×8 значение N должно лежать в пределах [3,4], так как обзора датчика хватит для определения не более 4 одновременно проходящих под ним человек. После каждого повтора максимуму и его соседям присваивается значение минимума, чтобы исключить повторную обработку одного и того же пятна.

Информация о принадлежности предыдущих пятен сохраняется. Для каждого из текущих определяется наличие у него «родителя» (пятна, которое было активно в том же столбце или в соседнем), если его нет, то создается запись для нового пятна. Далее для каждого из пятен обновляются значения *first*, *last* и *size* и определяется направление движения, аналогично алгоритму определения одного человека в кадре.

Заключение

В ходе работы была проведена оценка пропускной способности и энергоэффективности IPv6 over BLE для определения целесообразности обработки данных на оконечных устройствах, а также разработан алгоритм, позволяющий определять не только направление проходящих людей, но и с высокой точностью устанавливать их количество.

Результаты данной работы могут быть использованы при разработке недорогих и высокоточных решений для отслеживания проходимости помещений.

Список литературы

1. Petersen, H.(1), Wählich, M.(1) and Schmidt, T.C.(2) (2021) ‘Mind the gap: Multi-hop IPv6 over BLE in the IoT’, CoNEXT 2021 - Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies, pp. 382-396–396. doi:10.1145/3485983.3494847.

2. Gupta, N.K. (2016) Inside Bluetooth Low Energy, Second Edition. Norwood: Artech House. Available at:

<https://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=1511862&site=eds-live> (Accessed: 25 January 2023).

3. Robert Bosch GmbH (Stuttgart) (2021) ‘Measuring people-flow through doorways using easy-to-install IR array sensors’. Available at: <https://search.ebscohost.com/login.aspx?direct=true&db=edsprg&AN=edsprg.10948354&site=eds-live> (Accessed: 25 January 2023).

4. Perra, C. et al. (2021) ‘Monitoring Indoor People Presence in Buildings Using Low-Cost Infrared Sensor Array in Doorways’, Sensors (Basel, Switzerland), 21(12). doi:10.3390/s21124062.

ОБЗОР И АНАЛИЗ ТЕХНОЛОГИЙ МОНИТОРИНГА ПАРНИКОВЫХ ГАЗОВ

Кунижев И.Р.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе рассматриваются виды парниковых газов и их основные характеристики. Описывается взаимосвязь карбоновых полей и карбоновых единиц с эмиссией почвы. Исследованы методы оценки карбоновых единиц.

Введение

В настоящее время технологии развиваться ускоренными темпами, что приводит к появлению большого числа различных заводов и предприятий. Это влечет за собой увеличение выбросов различных газов в атмосферу, таких как углекислый газ, метан, озон и др. Для решения данной проблемы используются различные методы, например, энергосбережение [1]. Так же одним из способов является контроль выбросов с использованием карбоновых полигонов. Целью данной работы является анализ существующих методов оценки парниковых газов на карбоновом полигоне.

Задачами являются исследование характеристик парниковых газов, описание карбоновых единиц и их

связь с парниковыми газами, исследование способов количественной оценки парниковых газов при эмиссии почвы.

Парниковые газы и их основные характеристики

Согласно [2] парниковые газы – это газообразное составляющее атмосферы в видимом диапазоне, которое испускает и поглощает инфракрасное излучение. К таким газам относятся углекислый газ (CO_2), метан (CH_4), водяной пар (H_2O), озон (O_3). Они по своему происхождению подразделяются на природные и антропогенные, однако, например, водяной пар относится к обеим категориям.

При анализе воздействия парниковых газов на почву и атмосферу рассматривается количественная оценка выбросов и поглощения парниковых газов. Данной оценкой является концентрация газа в атмосфере или почве и измеряется в ppm (миллионная доля) или ppb (миллиардная доля). Так же количественную оценку можно фиксировать в процентах.

Одно из основных свойств парниковых газов связано с парниковым эффектом, а именно способность поглощать тепловую энергию [3], что позволяет защитить землю от радиационного и теплового излучения. Особенно большую роль в парниковом эффекте играет водяной пар.

При рассмотрении свойств отдельных газов можно так же выделить углекислых газ, концентрация которого в атмосфере растет с начала индустриальной эпохи. Именно данную тенденцию связывают с постепенным потеплением климата. Однако, стоит отметить, что диоксид углерода играет очень важную роль в процессах метаболизма живой клетки, но с другой стороны, при больших концентрациях он смертелен. Например, 15000 ppm является смертельной дозой для человека, в то время как его концентрация в атмосфере при нормальных условиях 400 ppm.

При изучении метана было установлено, что он наиболее устойчив к химическим воздействиям. Он является основным компонентом нефтяных, природных, болотных газов, а также его концентрация в атмометре составляет не более 2 ppm.

Озон является основным газом, защищающим от ультрафиолетовых лучей. Его концентрация в приземистых слоях Земли составляет 0.015 ppm, однако в стратосфере она составляет 2-8 ppm [4].

Карбоновые поля и карбоновая единица

Карбоновый (углеродный) полигон – это территориально распределенная научно-исследовательская площадка, где производится мониторинг выбросов и аккумуляции парниковых газов [5]. Создание таких площадок позволит выполнять 3 задачи:

- Расчет углеродного баланса местности с учетом эмиссии почвы с помощью специальной разработанной модели;

- Разработка метода отслеживания выбрасываемых и удерживаемых парниковых газов с использованием аэрокосмических данных;

- Разработка методов углеродного контроля

Размеры карбоновых полигонов могут быть разными, от единиц до скольких тысяч гектар; могут включать морскую площадку; рельеф может быть как равнинным, так и лесистым или болотистым.

Стоит отметить, что в 2020 г. Был запущен пилотный проект «Карбоновые полигоны» от Министерства обороны России, благодаря которому по всей стране были открыты исследовательские площадки и началось активное изучение параметров парниковых газов при эмиссии почвы. Так же данные исследования позволяют фиксировать количество углеродных единиц, которые имеются у той или иной местности.

Углеродная (карбоновая) единица – это квота на выброс углекислого газа в атмосферу. Согласно [6] у каждой компании есть определенные ограничения на выброс углекислого газа. Если предприятие производит выбросы меньше того ограничения, которое ему выделило государство, то оно получает квоту, которую можно передать другой компании, у которой выбросов больше. Таким образом происходит забота об экологии.

Методы оценки карбоновых единиц

Карбоновая единица измеряется в тоннах CO_2 -эквивалента. Одним из методов количественной оценки является расчетный метод, которые предполагает суммирование данных о содержании углерода в топливе, материалах, сырье. Которые использует конкретное предприятие.

С карбоновых полей данные собираются в миллионных долях. Далее происходит перевод ppm в mg/m^3 согласно уравнению Менделеева-Клапейрона:

$$p \cdot V = \frac{m}{M} \cdot R \cdot T \quad (1)$$

Полученные результаты переводят в т/Га. Таким образом, на выходе, получается количество углекислого газа в тоннах, которое выделяется при эмиссии почвы, пространством 1 гектар. Рассчитанная тонна и будет карбоновой единицей

К методом оценки выбросов можно отнести непрерывные измерения и периодические [7]. При непрерывных измерениях необходимо автономное оборудование, способное проводить периодические расчеты и отправлять сырые или обработанные данные в единую систему сбора данных. Так же они рассчитаны на крупные карбоновые полигоны.

Периодические измерения достаточны для небольших источников. В данном случае измерения осуществляются с помощью проб почвы.

Заключение

Существующие методы оценки парниковых газов позволяют в полной мере рассчитать количество углеродных единиц в карбоновом полигоне, однако для этого необходимо комплексное оборудование.

Список литературы

1. Белоусов В.Н., Смородин С.Н., Лакомкин В.Ю. Энергосбережение и выбросы парниковых газов (CO₂): учебное пособие/ СПбГТУРП. – СПб., 2014. – 52с. – ISBN 978-5-91646-071-1
2. ГОСТ Р ИСО 14064-1-2021. Газы парниковые. Часть 1. Требования и руководство по количественному определению и отчетности о выбросах и поглощении парниковых газов на уровне организации. – М.: Технический комитет по стандартизации ТК 20 «Экологический менеджмент и экономика», 2021
3. Голубев Г.Н Основы геоэкологии. Учебник для студентов высших учебных заведений. – М.: Изд-во ГЕОС, 1999. – 338с. – ISBN 5-89118-059-6
4. Кашкин В.Б. Стратосферный озон: вид с космической орбиты: монография/ В.Б. Кашкин, Т.В. Рублева, Р.Г. Хлебопрос. – Красноярск: Сиб. Федер. Ун-т, 2015. – 184с., ISBN 978-5-7638-3348-5
5. Абакумов Е.В., Поляков В.И., Чуков С.Н. Подходы и методы изучения органического вещества почв карбоновых полигонов России (обзор). // Почвоведение, 2022, №7, с.773-786
6. Российская Федерация. Законы. Федеральный закон об ограничении выбросов парниковых газов: Федеральный закон №296-ФЗ : [Принят государственной Думой 1 июня 2021 года: одобрен Советом Федерации 23 июня 2021 года]. – Москва: Кремль : Кодекс, 2021. – 20 с.
7. Астраханцева Е.А., Гатина Э.А. Углеродные единицы как новый вид инвестиций в АПК в условиях циркулярной экономики. // Russian journal of management, 2022, том 10 №1, с. 116-120

РАЗРАБОТКА АППАРАТНОЙ ЧАСТИ СТЕНДА ДЛЯ ПРОВЕДЕНИЯ МАССОВЫХ ЛАБОРАТОРНЫХ РАБОТ ПО ЭЛЕКТРОНИКЕ

Казакова В.Е., Хобов А.П.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Изложены текущие результаты разработки аппаратной части стенда для проведения массовых лабораторных работ по электронике.

Введение

Процедура проведения типовых лабораторных работ по курсу электроники включает в себя измерение характеристик электронных компонентов по точкам, с использованием обычных контрольно-измерительных приборов (источников сигнала, мультиметров). В процессе работы студенты знакомятся с особенностями работы различных типов компонентов, а кроме того, осваивают основные измерительные процедуры. В связи с тем, что лабораторное оборудование постепенно устаревает и выхо-

дит из строя, имеет смысл искать различные формы проведения работы с сохранением максимально возможной реалистичности процесса.

На рынке имеются готовые решения, однако они обладают заметными недостатками для данной задачи:

- реальные стенды от компаний labstand [1], профистенд [2], и др. — громоздкие, ненастраиваемые;
- VR-стенды — требуется дорогостоящее оборудование, производительный локальный компьютер;
- симуляторы: Tinkercad [3] — нет хороших моделей компонентов, Microcap [4] — отсутствует наглядный интерфейс и др.

В данной работе проблема решается за счёт разработки реального стенда с большим числом измеряемых компонентов и коммутирующей оснасткой, к которому осуществлялся бы удалённый доступ.

В отличие от аналогичных проектов, реализуемых в МИЭМ под руководством А. Ю. Романова, Д. А. Попова, А. Д. Касаткина, здесь не требуется сеансовый доступ к оборудованию на длительное время отдельным пользователям; для измерения отдельной точки характеристики требуется доступ к измерительному оборудованию на доли секунды. В этой связи очередь заданий должна иметь совсем другую конфигурацию: можно одновременно предоставлять доступ к большому числу измеряемых компонентов для большого числа пользователей по выданным вариантам.

В данной статье приведены текущие результаты разработки аппаратной части проекта.

1. Структура аппаратной части

Аппаратная часть проекта включает в себя стенд, комплект измерительных приборов, управляющий микроконтроллер. На стенде, который проектируется в форме набора взаимосвязанных печатных узлов, размещено большое число объектов измерения — электронных компонентов с различными элементами нагрузки, кроме того, средства подключения необходимых студенту объекта и нагрузки к измерительным приборам.

Объектами измерения в данном стенде являются диоды, биполярные и МОП-транзисторы. Некоторые модели, использованные для тестирования, приведены в списке ниже:

- КТ3102ГМ, Биполярный транзистор, NPN, 20В, 0.2А, 0.25Вт, 200МГц, h21е=400...800 [КТ-26 / ТО-92] (=BC548С)
- КП103К1, Транзистор, Р-канал, малой мощности [ТО-92 / КТ-26]
- 1N4007, Диод выпрямительный 1А 1000В [DO-41]

Всего в наборе измерительных объектов присутствуют 16 биполярных транзисторов, 16 полевых транзисторов и 16 диодов, которые могут быть своевременно заменены на новые благодаря монтажу компонентов на макетной плате. На основе данных компонентов студенты смогут снимать вольт-амперные характеристики элементов и сток-

затворные характеристики транзисторов по точкам, строить графики и проводить необходимые расчеты на их основе.

Для соединения компонентов с микроконтроллером и своевременной подачи питания на определенный объект измерения используется большое количество реле HJR1-2C-L-3VDC.

2. Разработка способа подключения микроконтроллера

Локальный микроконтроллер используется в проекте для управления процессом переключения измеряемых компонентов и нагрузок, а также для взаимодействия с контрольно-измерительными приборами: управления режимами их работы, считывания показаний и передачи их на управляющий компьютер.

Для решения поставленной задачи был выбран микроконтроллер STM32F103C8T6 на плате «Blue pill» от компании ST-Microelectronics: она обладает всем нужным функционалом, является доступной и для неё есть множество бесплатных и многофункциональных средств разработки.

Разработка прошивки производилась в STM32CubeIDE от разработчиков самого микроконтроллера, моделирование и тестирование функционала производилось в Proteus 8 [5].

Алгоритм действия микроконтроллера состоит в следующем:

1) микроконтроллер программируется на передачу и получение данных с управляющего ПК. Связь МК и компьютера реализована через шину USB, имеющуюся как на компьютере, так и на плате МК. Приём/передача команд осуществляется через terminal;

2) получив информацию о том, какие блоки на рабочей плате необходимо активировать, микроконтроллер задает каналы для вывода с них напряжения. Выбранный микроконтроллер работает от напряжения 3.3 В и выводит напряжение близкое к рабочему;

3) далее происходит замер нужных на макетной плате с помощью аналого-цифрового преобразователя;

4) измеренные данные передаются на ПК;

5) далее процесс может повторяться для других входных данных (от другого студента).

3. Результаты электрического моделирования измерительной оснастки

Для проверки правильности разработки электрической схемы измерения её работа была промоделирована с помощью средств САПР.

В проекте на данный момент спроектирована и реализована в программе моделирования Proteus схема подключения объектов измерения к микроконтроллеру, разработан алгоритм выбора конкрет-

ного элемента для данного измерения (в данном случае — биполярного pnp транзистора BC547) с помощью пары реле-коммутатор (рис. 1). Источники напряжения 3V имитируют напряжение, поступающее с микроконтроллера. Кнопка на реле имитирует подачу и прекращение подачи сигнала с микроконтроллера, что включает и выключает коммутатор. Три логических выхода, тоже подключенных к микроконтроллеру, регулируют работу коммутатора, переключая подачу напряжения на базу конкретного транзистора. Снятие выходной вольт-амперной характеристики будет производиться по точкам на коллекторе транзистора.

При моделировании множества каналов коммутатора важно проследить, чтобы ток утечки не был достаточно большим и не повлиял на точность значений. Т.к для каждого типа объектов измерения используется реле, коммутатор и восемь копий объекта, подключенных к коммутатору, необходимо обратить на это особое внимание. При подаче на включенный коммутатор напряжения от микроконтроллера, это напряжение подается на базу одного из транзисторов. В установившемся режиме необходимо сравнить ток активного транзистора и суммарный ток всех транзисторов, т.е ток источника (места подключения амперметров и показания измерений представлены на рис. 1).

По результатам моделирования получены следующие значения:

- ток активного транзистора — 2,895 мА
- суммарный ток всех транзисторов — 2,902 мА
- ток утечки — 7,1 мкА

Полученный ток утечки составляет 0,2% от общего тока, следовательно, этим фактором можно пренебречь.

4. Результаты разработки печатных узлов

На основе параметров, полученных при моделировании, была осуществлена трассировка печатного узла данного фрагмента схемы измерений в программе DipTrace [6]. Толщина дорожки рассчитана по максимальному току и равняется 0.3 мм. Корпуса, используемые при трассировке, соответствуют используемым элементам. Однако после печати платы выяснилось, что выбранная ширина дорожки слишком мала, и из-за такой ширины на плате возникают дефекты, которые приводят к неучтенным потерям тока и даже нарушениям прохождения тока по контактным дорожкам. Дальнейшая разработка печатной платы будет направлена на усовершенствование текущей трассировки, печать платы, подходящей по функциональным требованиям, установка компонентов и тестирование работы платы

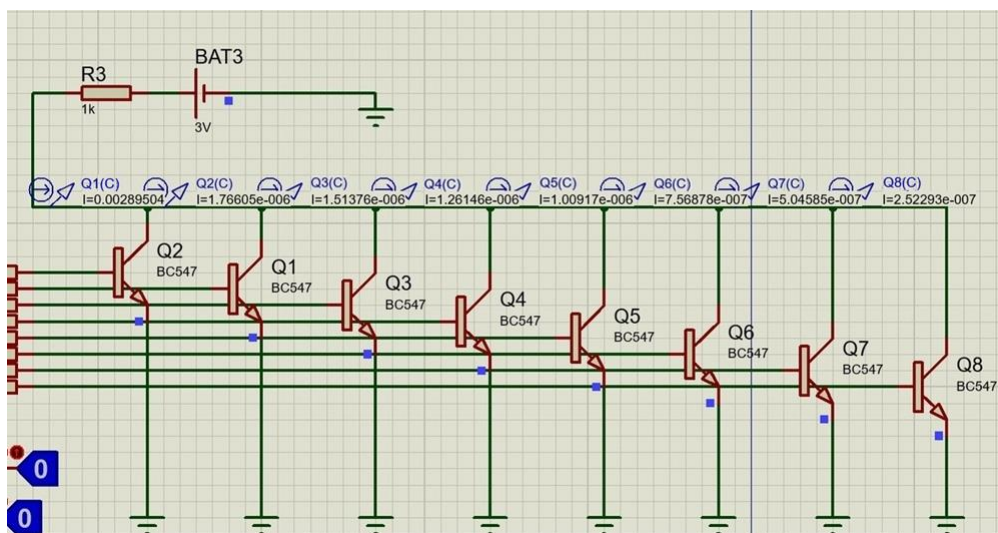


Рис. 1. Фрагмент системы для измерения характеристик компонентов одного типа (биполярных транзисторов)

Сетевое программное обеспечение

В данном проекте реализовано сетевое взаимодействие между виртуальной схемой на рабочем месте пользователя и физической, собранной в лаборатории. В клиентском приложении реализован графический интерфейс, позволяющий собрать виртуальную электрическую схему, используя предлагаемые виртуальные приборы, а также выполнить проверку корректности сборки данной схемы.

После верной сборки виртуальной схемы появляется возможность выполнять лабораторную работу согласно методическим указаниям (выставлять необходимые значения источника напряжения, снимать показания измерительных приборов). При изменении параметров схемы (напряжение и ток источника) клиентское приложение выполняет TCP-запрос к серверу. TCP-сервер, приняв входящее клиентское соединение, создает отдельный поток, в котором обрабатывает запрос. Из запроса извлекаются параметры настройки физической схемы, которые передаются на контроллер. Затем с контроллера сервер получает показания приборов и передает эти показания клиенту в ответ на запрос, после чего клиент закрывает сетевое соединение.

Полученные с сервера показания отображаются в клиентском приложении на соответствующих дисплеях виртуальных приборов.

Таким образом, сервер может обрабатывать несколько клиентских запросов, поступающих в одно время. Количество возможных одновременных подключений ограничивается ресурсами среды, в которой выполняется серверное приложение.

Заключение

Для решения ряда проблем, возникающих при проведении лабораторных работ по курсу электроники, было предложено решение, которое сочетает наглядность работы с исследуемыми объектами и

контрольно-измерительными приборами и массовость.

При проектировании аппаратной части системы был настроен управляющий микроконтроллер, проведено электрическое моделирование схемы измерений на примере биполярных транзисторов, подготовлен рисунок двухсторонней печатной платы.

В дальнейшей реализации измерительные компоненты можно будет заменять на требуемые в условиях лабораторной или по истечении срока их рабочего режима.

Рассмотрено сетевое взаимодействие программной и аппаратной части.

Список литературы

1. Типовой комплект учебного оборудования «Электроника», исполнение настольное, ручное ЭНП [Электронный ресурс] URL: <https://labstand.ru/catalog/electronika/tipovoj-komplekt-uchebnogo-oborudovaniya-elektronika-ispolnenie-nastolnoe-ruchnoe-e-nr> (Дата обращения: 28.01.2023)
2. Стенд "Электроника" ЭЛБ-241.053.01 [Электронный ресурс] URL: <https://profistend.info/catalog/stend-elektronika-elb-241-053-01.html> (Дата обращения: 28.01.2023)
3. AUTODESK Tinkercad [Электронный ресурс] URL: <https://www.tinkercad.com/dashboard> (Дата обращения: 28.01.2023)
4. SPECTRUM Software [Электронный ресурс] URL: <http://www.spectrum-soft.com/download/download.shtm> (Дата обращения: 28.01.2023)
5. Proteus [Электронный ресурс] URL: <https://www.labcenter.com/> (Дата обращения: 28.01.2023)
6. DipTrace [Электронный ресурс] URL: <https://diptrace.com/rus/> (Дата обращения: 28.01.2023)

РАЗРАБОТКА ВИРТУАЛЬНЫХ МОДЕЛЕЙ ИЗМЕРИТЕЛЬНЫХ ПРЕОБРАЗОВАТЕЛЕЙ

Сергеенко К.А., Дубильер Я.А., Санникова Н.В.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В ходе данной работы были разработаны виртуальные лабораторные стенды с моделями аналоговых и цифровых датчиков в среде графического программирования NI LabVIEW. Такая разработка позволит проводить лабораторные работы по измерительным технологиям в дистанционном формате, а также подготовить студентов к использованию реальных датчиков и плат сбора данных.

Введение

Оснащение учебных лабораторий современными измерительными приборами и оборудованием обходится крайне дорого, а использование приборов возможно только в непосредственном присутствии студентов в лаборатории. Соответственно, создание и использование виртуальных стендов по сей день является актуальной задачей для выполнения фронтальных и дистанционных лабораторных работ.

Также увеличение интереса людей к профессиональному саморазвитию приводит нас к росту популярности онлайн-образования. В связи с этим появляется потребность в разработке специализированных онлайн-программ, благодаря которым любой желающий может повысить свою квалификацию.

Целью данной работы является обеспечение дистанционного выполнения лабораторных работ по измерительным технологиям, увеличение числа рабочих мест и снижение стоимости оснащения лаборатории.

В рамках работы решаются задачи по разработке нескольких виртуальных лабораторных стендов с датчиками для измерения температуры, влажности, расстояния. В стендах используются как цифровые, так и аналоговые датчики, также в работу виртуальных стендов внедрена модель платы сбора данных Arduino[1].

Одна из статей [2] рассматривает вопрос внедрения виртуальных лабораторных стендов в образовательный процесс, указывая на необходимость проведения виртуальных лабораторных работ в качестве демонстрационных или тренировочных, непосредственно перед реальным практикумом. Также в ряде исследований [3,4], проведенных в образовательных

учреждениях Китая, рассматривался вопрос проведения лабораторных работ, путем организации удаленного доступа к реальным устройствам. Авторы обеих работ разработали полноценную онлайн-платформу с 4 блоками виртуальных экспериментов, включающими около 40 исследований.

В современной образовательной литературе уже есть примеры созданных виртуальных лабораторных практикумов с измерительными преобразователями российской разработки [5]. Данные лабораторные стенды хорошо зарекомендовали себя в учебном процессе. Однако, их функционал и гибкость несколько ограничены, и они не предоставляют возможность студентам полностью ознакомиться с внешним видом и функциями реальных преобразователей, также в данных виртуальных стендах используются только аналоговые датчики, которые подключаются к приборам, и нет работ с использованием плат сбора данных.

Анализ существующих датчиков

В рамках подготовительной работы был проведен анализ нескольких аналоговых и цифровых датчиков различных физических величин.

Были проанализированы датчики для измерения следующих величин: температура, влажность, упругая деформация, давление, расстояние, освещенность, величина магнитного поля, ионизирующее излучение.

На рис. 1 представлена диаграмма, показывающая результаты анализа. На схеме можно увидеть найденные физические величины, которые можно измерить с помощью датчиков, и, непосредственно, сами разновидности датчиков.

После проведения исследования были выбраны 4 величины, датчики для измерения которых наиболее часто используются на практике, а именно:

- 1) Температура (выбрана аналоговая термопара для мультиметра и цифровой датчик температуры DH11);
- 2) Влажность воздуха и почвы (выбран аналоговый емкостной датчик влажности почвы и цифровой датчик влажности DH11);
- 3) Расстояние (выбран цифровой ультразвуковой датчик расстояния HC-SR04).
- 4) Угол поворота (выбран инкрементальный энкодер).
- 5) Давление (выбран цифровой датчик давления BMP180).

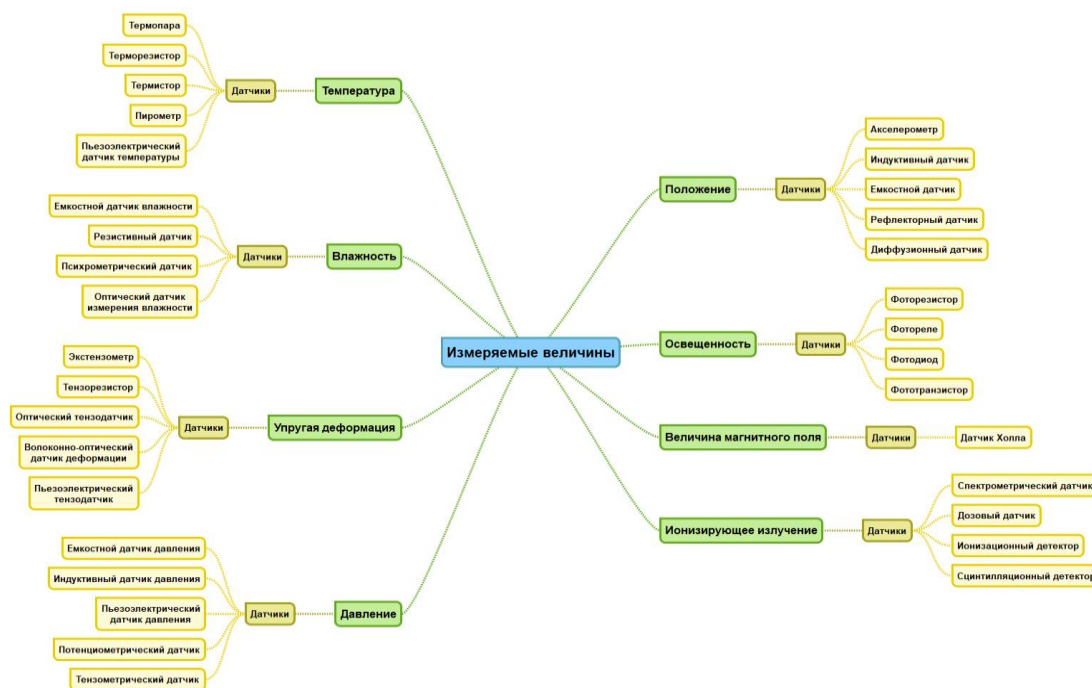


Рис. 1. Mindmap с результатами анализа

Лабораторные работы по измерению температуры, влажности и расстояния

Измерение выбранных величин температуры, влажности и расстояния осуществляются с помощью виртуальных моделей мультиметра и платы Arduino Uno. В случае аналогового датчика лабораторный эксперимент проходит с моделью мультиметра UT33D, измерения с цифровым датчиком осуществляются с Arduino Uno. Пользователь осуществляет виртуальное подключение датчика с дополнительными элементами для лабораторного эксперимента и платой Arduino Uno или мультиметром.

Для лабораторного эксперимента по измерению влажности используется горшок с цветком и стакан с водой, для измерения температуры - термокамера. В качестве имитации измерения расстояния используется автомобиль со стенами.

Передние панели всех виртуальных лабораторных стендов изображены на рисунке 2.

Фрагмент блок-диаграммы с кодом для виртуального лабораторного стенда по измерению влажности почвы можно увидеть на рисунке 3.

Все элементы в лабораторном стенде взаимодействуют между собой с помощью следующих глобальных переменных:

- «Global connectors.vi» – глобальная переменная, в которую поступает информация (в виде массива) о подключенных контактах и их номерах, а также о режиме работы стенда.
- «Global variable.vi» - глобальная переменная для передачи значений измеренных физических величин на мультиметр.
- «Values.vi» - глобальная переменная для передачи измеренных значений для платы Arduino.

Правильное подключение элементов и приборов осуществляется с помощью программы контроля подключений «Connectors check.vi», которая сравнивает массив с номерами подключенных контактов и эталонный массив со схемой верного подключения приборов.

Виртуальные лабораторные стенды запускаются из файла формата .exe, который не требует установки среды NI LabVIEW. В дальнейшем планируется интеграция в “Виртуальную лабораторию” по электро-радиоизмерениям [6], а также в рамках обучающих курсов для онлайн образования.

Внедрение виртуальных моделей измерительных преобразователей в учебный процесс

Разработанные лабораторные стенды позволяют учащимся НИУ ВШЭ, а также студентам других учебных заведений, участвовать в процессе обучения без вынужденного присутствия в лабораториях. С помощью методических материалов студенты смогут реализовать в режиме онлайн аналогичный процесс работы реальных датчиков на основе виртуальных приборов. Использование виртуальных стендов направлено на то, чтобы преподаватель без дополнительных внешних ресурсов смог продемонстрировать своим учащимся весь процесс функционирования работы датчиков в режиме реального времени. Также виртуальная разработка позволит студентам иметь доступ к нужным материалам в любое доступное время для самостоятельного изучения или закрепления полученных знаний, без участия преподавателя и необходимого доступа в учебные аудитории.

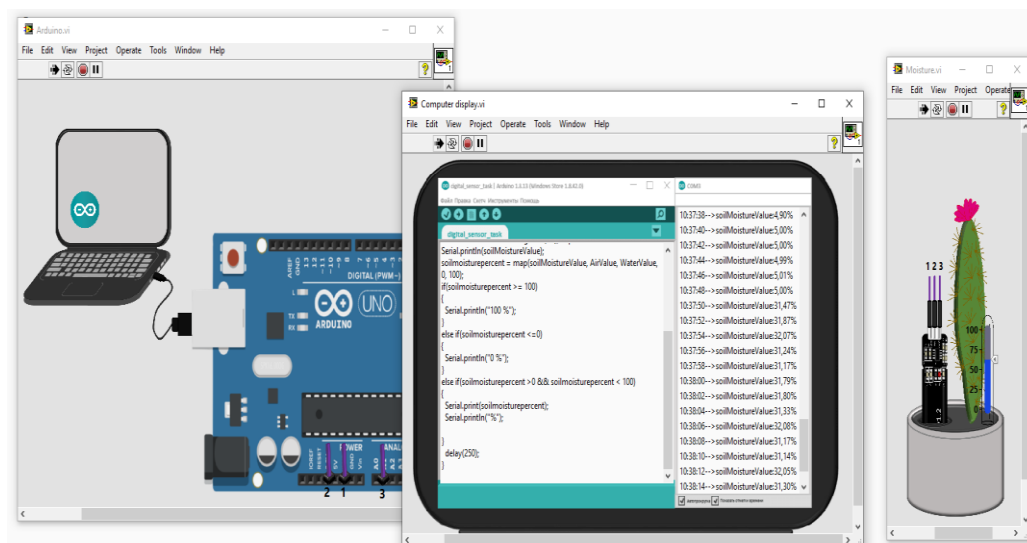


Рис. 2. Виртуальный лабораторный стенд для измерения влажности почвы

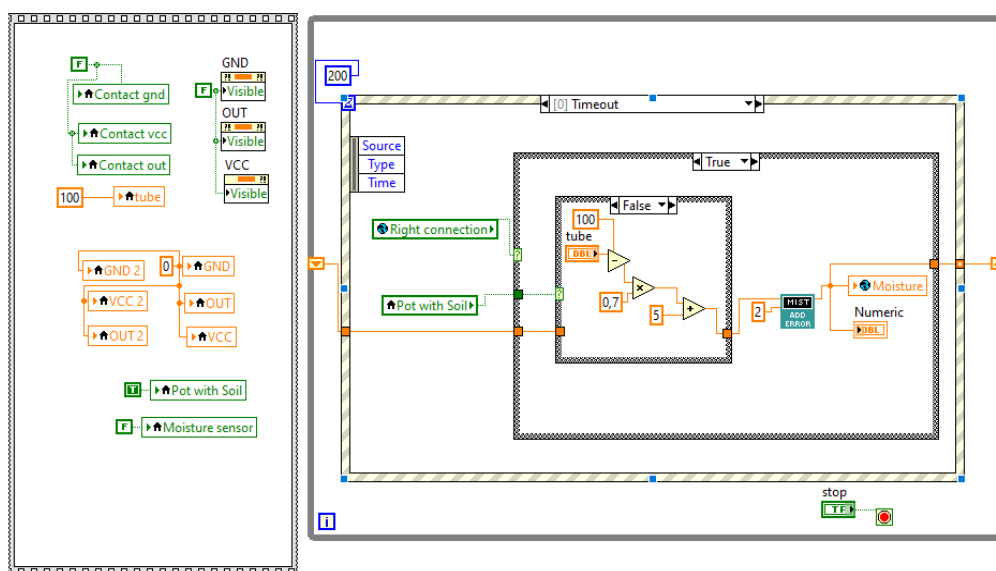


Рис. 3. Блок-диаграмма ВП "Moisture.vi"

Заключение

В завершение отметим, что разработка виртуальных стендов не только позволит предоставить возможность студенту обучаться в онлайн-режиме, но и также избежать затраты на нужное или дополнительное оборудование. В перспективе такой подход можно реализовать в комплексный учебный онлайн-курс, который будет доступен на открытых образовательных платформах любому желающему.

Список литературы

1. Arduino.ru: сайт. – 2021. – URL: <http://arduino.ru/> (Дата обращения: 25.10.2021).
2. Юрий А. И., Красивская М. И., Дмитриев А. В., Злодеев Г. Ю. Применение виртуальных лабораторных стендов в образовательном процессе // Информационные технологии. 2014. № 6. С. 70-72.

3. Yongyu Peng, Kehong Zhang. Based on LabVIEW Remote Virtual Electronic Laboratory Design and Implementation. -The Open Cybernetics & Systemics Journal, 2014.

4. Yi Z., Jian-Jun J., Shao-Chun F. A LabVIEW-based, interactive virtual laboratory for electronic engineering education //International Journal of Engineering Education. – 2005. – Т. 21. – №. 1. – С. 94-102.

5. Батоврин В. К. и др. LABVIEW. Практикум по основам измерительных технологий. Учебное пособие. – ДМК Пресс, 2009.

6. Красивская М.И., Юрий А.И., Гаспарян А.О. и др. Виртуальная лаборатория по электро-радиоизмерениям // Инженерное образование. 2021. № 30. С. 67-76.

ИССЛЕДОВАНИЕ И РАЗРАБОТКА МЕТОДИКИ РАСЧЕТА МОНИТОРИНГА ПОЛЯ, СОЗДАВАЕМОГО КОНИЧЕСКОЙ РУПОРНОЙ АНТЕННОЙ

Кучина А.П.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе рассматривается методология по расчету элементов поля, создаваемых коническим рупором. Кроме того, удалось уточнить математическую модель для определения плотности потока энергии вблизи апертуры конической рупорной антенны. В исследовании получены результаты построения диаграммы направленности в Mathcad в зависимости от фазировки.

Введение

На протяжении многих лет передача и прием информации с помощью апертурных антенн является важной областью, требующей постоянного развития. Рупорная антенна является, пожалуй, самым популярным типом апертурных антенн. Рупорные антенны широко используются для передачи и приема электромагнитных волн в таких областях, как микроволновая связь, тестирование электромагнитной совместимости. Также они нашли свое применение в качестве силовых элементов для радиоастрономии, спутникового слежения и антенн связи [1].

Существуют различные типы рупорных антенн: секторный рупор (Е-плоскость и Н-плоскость), пирамидальный рупор и коническая рупорная антенна. Имеется тенденция, что в основном рассматриваются характеристики излучения в Е- и Н-плоскостях секторных и пирамидальных рупоров, поэтому их характеристики излучения хорошо проанализированы, обобщены и получены диаграммы направленности [2]. Однако диаграмма направленности конической рупорной антенны, несмотря на ее популярность и широкий спектр применения, не получила должного внимания по сравнению с другими из-за сложности расчета параметров излучения.

Коническая рупорная антенна обеспечивает достаточно высокий коэффициент усиления и выдает большую мощность. Преимуществами конических рупорных антенн является возможность работы с чрезвычайно хорошими характеристиками направленности и коэффициента усиления, также неоспоримым преимуществом этой антенны является низкий уровень заднего лепестка. Предполагается, что характеристики излучения конического рупора сильно зависят от распределения амплитуды и фазы по апертуре рупора [3]. В настоящее время общепризнано, что при измерении мощности излучения конической рупорной антенны можно использовать метод интегрального дальнего поля, так что мощность излучения получается путем интегрирования плотности мощности поля излучения в пределах определенной поверхности пространства. Как правило, распределение

плотности мощности на поверхности измеряется в процессе практической эксплуатации.

Необходимость проведения исследования на выбранную тему связана с тем, что в результате анализа аналогичных работ по этой теме необходимы дополнительные исследования для расчета коэффициента направленности в зоне Френеля, также учитывающие фазировку рупорной апертуры [4]. Впоследствии более точные результаты могут быть использованы в моделях, которые прогнозируют электромагнитную обстановку.

В данном исследовании предлагается новый метод получения диаграммы направленности конической рупорной антенны, который решает проблему получения более точных характеристик излучения.

Целью данного исследования является изучение параметров конической рупорной антенны, получение диаграммы направленности ее излучения методом интегрирования с учетом фазы апертуры рупора.

Теория волноводов и рупорных антенн

Антенны в форме открытого волновода обычно слабо направлены и требуют большой излучающей поверхности для достижения остронаправленных характеристик. Однако размеры отверстия волновода не могут быть слишком большими, иначе волны высшего порядка могут нарушить нормальную работу волновода. Кроме того, волноводные излучатели плохо согласуются со свободным пространством из-за большого коэффициента отражения, вызванного резким изменением распространения электромагнитных волн. Рупорные излучатели преодолевают эти недостатки путем плавного увеличения поперечного сечения волновода, что снижает интенсивность волн высшего порядка и улучшает координацию со свободным пространством. Конический рупор обычно возбуждается участком круглого волновода, в котором распространяется волна типа H_{11} , и имеет неустойчивую плоскость поляризации, создавая в рупоре сферическую волну с аналогичным распределением амплитуды и фазы, что и в круглом волноводе.

Методы анализа ДН конических рупоров

Анализ геометрической теории дифракции, приведенный в [5], — это метод, позволяющий точно предсказать диаграммы направленности (ДН) излучения конических рупоров в Е- и Н-плоскостях с использованием теорий Куомджяна и Патхака и метода наклонной дифракции. Этот метод может точно рассчитать главный луч, ближние и дальние боковые лепестки и задние лепестки. Кроме того, можно рассчитать диаграммы излучения широкополосных гофрированных конических рупоров в главном луче, который содержит большую часть излучаемой энергии (–40 дБ относительно поля бокового обзора) более эффективным способом по сравнению с традиционными методами, такими как интегрирование Кирхгофа или метод расширения сферической волны. Однако данный метод по расчету коэффициента направленного действия (КНД) конического рупора не учитывает фазировку апертуры, что приводит к значительному завышению расчетных значений КНД.

Важнейшим аспектом решения вопросов, связанных с оценкой электромагнитной безопасности, является определение плотности потока энергии электромагнитного излучения [6]. Это, в свою очередь, требует в качестве предварительной подготовки расчета таких параметров излучателя как нормированная характеристика направленности и коэффициент направленного действия. Настоящая методика описывает процесс расчета перечисленных параметров и плотности потока энергии для конического рупора с прямолинейной формой образующей, как показано на рис. 1.

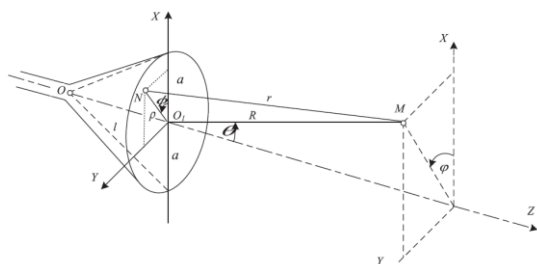


Рис. 1. Пояснение к методу расчета ДН конической рупорной антенны

Возбуждение круглого волновода производится волной типа H_{11} . Форма рупора определяется такими параметрами как длина рупора (l) и радиус апертуры (a). Предполагается, что плоскость, содержащая вектор напряженности электрического поля E , соответствует углу $\varphi = \pi/2$. Для зоны Френеля поле в точке наблюдения можно представить в следующем виде:

$$E(\theta, \varphi) = -i \frac{E_0 a^2}{2R\lambda} (1 + \cos \theta) e^{-ikR} f(\theta, \varphi), \quad (1)$$

где

$$f(\theta, \varphi) = \int_0^{2\pi} \int_0^1 f(\rho, \phi) e^{-iY\rho^2 - i\delta\rho \cos(\varphi - \phi)} \rho d\rho d\phi. \quad (2)$$

Далее приведена зависимость, которая аппроксимирует амплитудное распределение плоскости раскрыва рупора, возбужденного при помощи волны H_{11} :

$$f(\rho, \phi) = (1 - 0,37\rho^2 + (-0,845 + 0,215\rho^2))\rho^2 \cos^2 \phi. \quad (3)$$

Параметр обобщения

$$\delta = \frac{2\pi}{\lambda} a \sin \theta. \quad (4)$$

Для учета результирующей расфазировки используем:

$$Y = \frac{ka^2}{2l} + \frac{ka^2}{2R}. \quad (5)$$

Далее поясним (5). Член $\frac{ka^2}{2l}$ рассматривает квадратичную фазировку, которая возникает вследствие разницы в длинах путей текущей точки $N(\rho, \phi)$ в апертуре и точки O . Член $\frac{ka^2}{2R}$ принимает во внимание квадратичную фазировку, учитывающую разность

хода от наблюдаемой точки $M(R, \theta)$ и точек N (в апертуре) и O (в центре).

Результаты исследования

Далее показана нормированная диаграмма направленности (ДН) конической рупорной антенны (рис.2) для E -плоскости ($\phi = \pi/2$). Формулы (2) – (5) были взяты в основу расчета характеристик, в качестве аргумента была взята обобщенная координата для значений Y квадратичной фазировки в диапазоне от π до 4π .

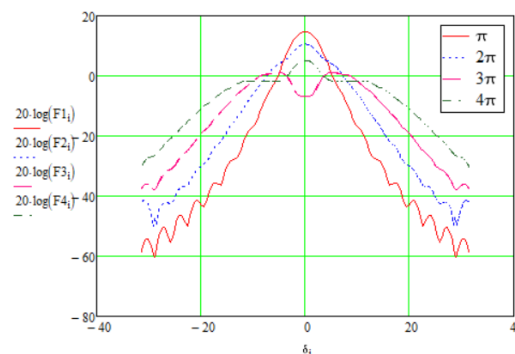


Рис. 2. Расфазовка диаграммы направленности в плоскости E конической рупорной антенны

Заключение

В результате произведенного улучшения математической модели, используемой для расчета значений диаграммы направленности и КНД вблизи апертуры конической рупорной антенны, можно сказать о следующей практической значимости: данная методика поможет более точно производить измерения электромагнитного излучения в местах размещения радиоприборов, тем самым обезопасив людей от облучения.

Список литературы

1. A. Bala and G. Saini. Optimization of Horn Antenna using various Techniques: A Review. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 5, no. 10, pp. 7806-7809, 2016.
2. H. Kobayashi. Horn Antenna. Advances in Environmental Engineering and Green Technologies, pp. 144-177, 2020.
3. N. A. Aboserwal, C. A. Balanis and C. R. Birtcher. "Conical Horn: Gain and Amplitude Patterns," in IEEE Transactions on Antennas and Propagation, vol. 61, no. 7, pp. 3427-3433, July 2013.
4. B. Polat and R. Daşbaşı. "Analysis of Conical Horn Antenna Radiation over Ground by Directional Currents Method," 2021 20th International Conference on Microwave Techniques (COMITE), 2021, pp. 1-6.
5. M. Narasimhan and M. Sheshadri. "GTD analysis of the radiation patterns of conical horns," in IEEE Transactions on Antennas and Propagation, vol. 26, no. 6, pp. 774-778, November 1978.
6. V. Kubanov, V. Ruzhnikov, and M. Spodobayev. "Electrodynamic model of electromagnetic fields radiation of conical horn antenna for purposes of electromagnetic safety." Infokommunikacionnye tehnologii, vol. 15, no. 1, pp. 97-102, 2017.

СОЗДАНИЕ АМОΡФНЫХ И АМОΡФНО-НАНОКРИСТАЛЛИЧЕСКИХ МЕТАЛЛИЧЕСКИХ СПЛАВОВ С ЗАДАННЫМИ СЛУЖЕБНЫМИ СВОЙСТВАМИ

*Саповский И.М.,
Рахматуллаев Т.Р., Пинчук К.Е.
Дальневосточный федеральный университет,
департамент общей и
экспериментальной физики*

Аннотация

В статье рассматриваются вопросы, связанные с разработкой функциональных и конструкционных аморфных и аморфно-нанокристаллических металлических сплавов, обладающих уникальными свойствами. Авторы используют метод спиннингования для получения быстрозакаленных сплавов на основе железа (Fe-Co-Ni-Si-B), и представляют результаты исследования структуры, свойств и процессов формирования данных материалов, что актуально для создания новых материалов с необходимым комплексом служебных свойств.

Введение

Научно-технологический прогресс в развитии современной науки и техники постоянно выдвигает все новые требования к разрабатываемым и создаваемым материалам. Одним из приоритетных направлений наук о материалах является разработка функциональных и конструкционных аморфных и аморфно-нанокристаллических металлических сплавов, обладающих уникальным сочетанием физико-химических и механических свойств. Важной задачей по совершенствованию существующих и созданию новых аморфных материалов является определение физических закономерностей их получения и определение связи структура - свойства.

Таким образом, для создания новых материалов с требуемым комплексом служебных свойств актуальным является изучение и понимание физической природы процессов, определяющих формирование всего спектра пространственных неоднородностей в аморфных металлических средах в зависимости от технологии получения. В последнее время широкое распространение для получения аморфных сплавов [1-5] приобрел метод быстрой закалки из расплава - метод спиннингования.

Целью данной работы является изучение процесса получения быстрозакаленных сплавов на основе железа методом спиннингования и исследование свойств полученных сплавов.

Материалы и методы

Быстрозакаленные сплавы на основе железа (Fe-Co-Ni-Si-B) были получены с использованием установки Melt Spinner SC. Данная установка включает в себя высоковакуумную систему, которая имеет вращающийся с большой скоростью (до 60 м/с) медный барабан, охлаждаемый с помощью водяного насоса. Диаметр медного колеса составляет 210 мм. Закалка

расплава в установке Melt Spinner SC проходит в атмосфере аргона (вакуум до 5×10^{-5} мбар). Сценарии перехода спиннингованных лент в равновесное состояние исследовались с использованием дифференциальной сканирующей калориметрии (Netzsch DSC 404 F1Pegasus). Анализ структуры поверхностей раздела образцов проводился методом электронной микроскопии (растровый электронный микроскоп Carl Zeiss Ultra 55+ и двухлучевая система Carl Zeiss Crossbeam 1540XB). Структура быстрозакаленных сплавов на атомном уровне контролировалась с помощью рентгеновской дифрактометрии. Изменение относительного электросопротивления спиннингованных лент от температуры фиксировалось с использованием четырех-зондовой установки с линейным расположением зондов.

В таблице 1 указаны параметры спиннингования сплава $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$: v / V – частота/скорость вращения барабана-холодильника; t – толщина ленты; H – ширина ленты. Для исследования в данной работе были выбраны образцы (плавки № 6, 7, 17), полученные при разных скоростях спиннингования, таблица 1. Все другие параметры быстрой закалки были постоянны: температура расплава $T_p = 1430^\circ\text{C}$; зазор между тиглем и барабаном $h = 0,15$ мм; давление в камере $P = 50 \pm 5$ мбар; разница давлений в камере и газовой магистрали $\Delta P = 200 \pm 5$ мбар. Выбор температуры расплава 1430°C определялся задачей получения аморфного состояния и необходимого значения коэффициента вязкости жидкости перед закалкой. Для этого температура T_p должна быть выше $T_{пл}$ не менее, чем на 200°C (для $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$ $T_{пл} \sim 1200^\circ\text{C}$).

Таблица 1. Параметры быстрой закалки расплава Fe-Co-Ni-Si-B

№ плавки	v (Гц) / V (м/с)	t (мкм)	H (мм)
6	40/25,1 \pm 0,1	22,9 \pm 1	10
7	45/28,3 \pm 0,1	16,6 \pm 0,5	10
17	50/37,7 \pm 0,1	55 \pm 1	5

Результаты

Исследуемые в данной работе быстрозакаленные магнитомягкие ($H_c \sim 0,02$ Э) сплавы Fe-Co-Ni-Si-B в исходном состоянии рентгеноаморфны. На рисунке 1 приведены термограммы спиннингованных лент $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$, полученные на дифференциальном сканирующем калориметре со скоростью нагрева 10 град/мин. Структурная релаксация сплавов $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$ в пределах аморфного состояния протекает до $T \sim 540^\circ\text{C}$ и сопровождается уменьшением свободного объема, снятием напряжений прокатки, формированием прекурсора перед кристаллизацией.

Отметим, энергия, выделенная при преобразовании структуры в пределах аморфного состояния (W_{as}) разная для образцов плавки 6, 7 и 17, таблица 2.

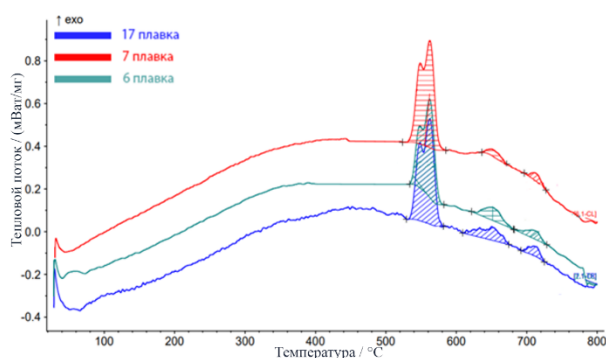


Рис. 1. Термограммы образцов $\text{Co}_{58}\text{Ni}_{10}\text{Fe}_5\text{B}_{16}\text{Si}_{11}$ для плавки 6, 7, 17

Для ленты плавки 7, закаленной при большей скорости спиннингования, относительно плавки 6, $W_{as} = 769.1$ Дж/г, поэтому процент аморфной фазы, выше [2-4].

Таблица 2. Основные характеристики термокривых образцов $\text{Co}_{58}\text{Ni}_{10}\text{Fe}_5\text{B}_{16}\text{Si}_{11}$

Образец	ν , Гц	W_{0-1} , Дж/г	W_2 , Дж/г	W_3 , Дж/г	W_{as} , Дж/г
6	40	61.4	10.6	5.6	701.9
7	45	71.7	4.1	3.3	769.1
17	50	70.1	12.9	5.1	647.8

Увеличение частоты вращения барабана-холодильника до 50 Гц (плавка 17) сопровождается уменьшением W_{as} (647,8 Дж/г). Анализ фрактограмм образца 17 [6] показал, что уже в исходном состоянии данный быстрозакаленный сплав является аморфно-нанокристаллическим, что приводит к меньшим значениям W_{as} .

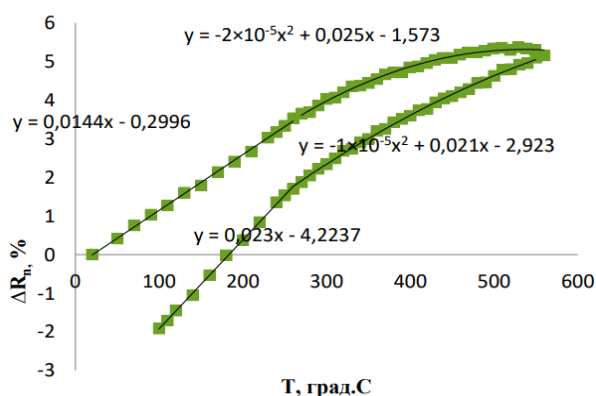


Рис. 2. Зависимость удельного сопротивления от температуры для образца плавки № 7

Данный вывод подтверждается, характером изменения удельного электрического сопротивления (рис. 2): при нагреве до $T = 560$ °C и последующем охлаждении наблюдается гистерезис с $\Delta R < 0$. Наличие двух участков на графике $\Delta R = f(T)$ свидетельствует

о смене механизма структурной релаксации. Низкотемпературный отжиг сопровождается выходом свободного объема, далее происходит снятие напряжений прокатки [7]. Многоэтапный процесс перехода быстрозакаленных сплавов $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$ в равновесное состояние при температуре выше 540°C характеризует второй участок термокривой, рис. 1. Вне зависимости от скорости спиннингования на термограммах наблюдается три экзотермических пика, характеристики которых приведены в таблице 2, причем первый пик – раздвоенный (W_{0-1} – приведенные энергии соответствующих пиков). Следовательно, в результате кристаллизации аморфного сплава выделяется четыре фазы, и две из них имеют близкие температуры кристаллизации.

На рисунке 3 приведен профиль рентгенограммы быстрозакаленного сплава $\text{Co}_{58}\text{Ni}_{10}\text{Fe}_5\text{B}_{16}\text{Si}_{11}$, плавка 6, который подтверждает переход в аморфно-нанокристаллическое состояние: на аморфное гало наложены дополнительные «кристаллические» пики, которые свидетельствуют о наличии нанокристаллов в аморфном материале.

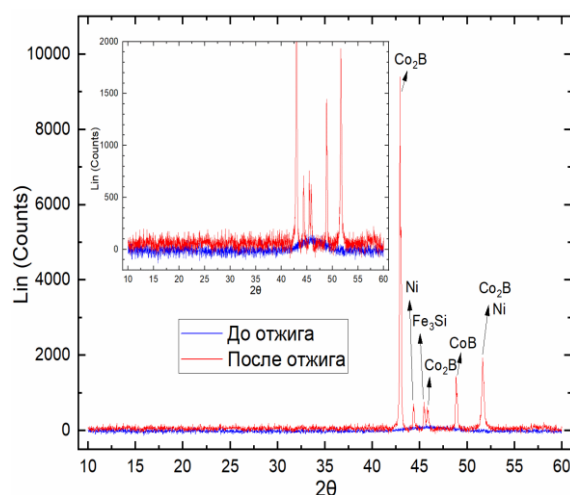


Рис. 3. Профиль рентгенограмм для образца плавки №6 до и после отжига

Использование метода спиннингования при получении аморфных сплавов предполагает наличие у них двух поверхностей раздела – контактной к медному барабану-холодильнику и свободную от его непосредственного влияния. На рисунке 4 а, в, д представлены электронно-микроскопические изображения контактной поверхности исследованных спиннингованных лент сплава Co-Ni-Fe-Si-B, характеризующееся развитым рельефом в виде полос прокатки и каверн различной формы и размера. Спектральный Фурье – анализ морфологии поверхностей раздела, рис. 4, показывает, что наименьшей степенью анизотропии неоднородностей на контактной поверхности обладает образец плавки №6 ($V = 25,1$ м/с). Увеличение скорости вращения барабана отражается на изменении формы каверн – от вытянутых вдоль оси прокатки тетраэдров (плавка №7, $V = 28,3$ м/с) до протяженных воздушных карманов (плавка 17, $V = 37,7$ м/с).

Проведенное аналогичное исследование свободной поверхности, рис. 4 б, г, е спиннингованных лент $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$, которая характеризуется более ровным бугорковым рельефом, показало, что прокатка оказывает опосредованное влияние на структуру данной поверхности, что отражается в анизотропных спектральных характеристиках (рис. 4 б, г, е). Следует отметить совпадение вида спектральных характеристик обеих поверхностей для режима $V = 28,3$ м/с, (рис. 4 в и г). Степень анизотропии свободной поверхности для быстрозакаленного сплава $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$ при $V = 25,1$ м/с выше, чем для контактной, что связано с анизотропно сформированным бугорковым рельефом, рис. 4 г.

Абсолютно разными характеристиками распределения неоднородностей по поверхностям раздела обладает быстрозакаленный сплав $\text{Fe}_5\text{Co}_{58}\text{Ni}_{10}\text{Si}_{16}\text{B}_{11}$, полученный при скорости вращения барабана $V = 37,7$ м/с: от высоко анизотропного ансамбля дефектов на контактной стороне (рис. 4 д) до полной изотропии свободной поверхности (рис. 4 е).

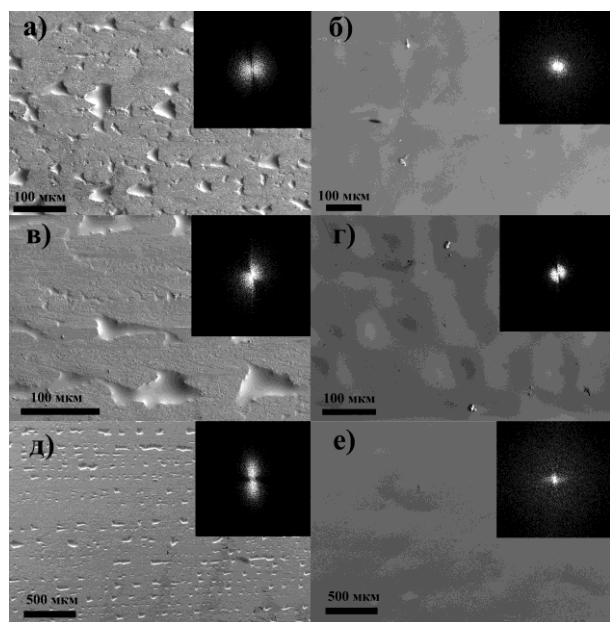


Рис. 4. Электронно – микроскопические изображения контактной и свободной поверхностей спиннингованных лент сплава Co-Ni-Fe-Si-B , полученных при разных скоростях вращения медного барабана: а, б - плавка 6; в, г - плавка 7; д, е – плавка 17

Таким образом, для аморфных сплавов (образцы 6, 7) технология спиннингования одинаковым образом влияет на формирование рельефа обеих поверхностей раздела. Наиболее оптимальным режимом с точки зрения формирования аморфной структуры поверхностей раздела с близкими характеристиками является режим $V = 28,3$ м/с (~ 30 м/с).

В зависимости от режимов получения были исследованы торцы спиннингованных лент $\text{Co}_{58}\text{Ni}_{10}\text{Fe}_5\text{Si}_{11}\text{B}_{16}$. Наличие градиента температур по толщине расплава от закалочного диска до свобод-

ной поверхности приводит к сложному фронту затвердевания материала, что и отражается в иерархической структуре торцов быстрозакаленных сплавов.

Заключение

В результате проведенных исследований по быстрой закалке из расплава сплавов системы Fe-Co-Ni-Si-B сформулированы условия получения аморфных и аморфно-кристаллических сплавов. Оптимальным режимом с точки зрения формирования морфоструктуры поверхностей раздела с близкими характеристиками является режим $V = 28,3$ м/с; для которого характерно наличие периодичностей минимального размера при высокой степени изотропности их распределения на контактной и свободной сторонах.

Изучена эволюция структуры аморфных быстрозакаленных сплавов $\text{Co}_{58}\text{Ni}_{10}\text{Fe}_5\text{B}_{16}\text{Si}_{11}$, которая сопровождается выделением энергии в связи со снятием напряжений прокатки, выходом свободного объема, с последующим поглощением энергии из-за подготовки прекурсорного состояния перед кристаллизацией. Рост скорости нагрева лент сопровождается увеличением выделенной энергии как в аморфном состоянии W_{as} , так и при кристаллизации.

Финансирование

Работа выполнена при финансовой поддержке Фонда ДВФУ, проект №22-02-03-005.

Список литературы

1. Филонов, М.Р. Теоретические основы производства аморфных и нанокристаллических сплавов методом сверхбыстрой закалки / М.Р. Филонов, Ю.А. Аникин, Ю.Б. Левин – Москва.: Изд-во МИСиС – 2006 – С. 326.
2. Кекало, И. Б. Процессы структурной релаксации и физические свойства аморфных сплавов, в 2 т. / И. Б. Кекало; Москва: изд-во дом МИСиС, 2016. – Т.1. – 436 с., Т.2. – 834 с.
3. Судзуки, К. Аморфные металлы. / К. Судзуки, Х. Фудзимори, К. Хасимото; под ред. Ц. Масумото – Москва: Металлургия, 1987. – 328 с.
4. Хандрих, К. Аморфные ферро- и ферримагнетики / К. Хандрих, С. Кобе; пер. с немецкого – Москва: Мир, 1982 – 296 с.
5. Глезер, А.М. Нанокристаллы, закаленные из расплава. / И.Е. Пермякова // – Москва: ФИЗМАТЛИТ, 2012. – 360 с.
6. Ткачев В. В. Эволюция структурных неоднородностей аморфных и аморфно-нанокристаллических сплавов системы $\text{Fe-(Cu, Nb)-(Si, B)}$ в процессах структурной релаксации / Автореферат дисс. на соискание ст канд. наук, 2019 – 22 с.
7. A. Kotvitckii, G. Krainova, A. Frolov, V. Ivanov, V. Plotnikov. Structure evolution of Fe- and Co-based amorphous alloys studied by electrical resistivity measurements / Solid State Phenomena Vol. 215 (2014) pp.185-189.
doi:10.4028/www.scientific.net/SSP.215.185.

ИССЛЕДОВАНИЕ ОПТИЧЕСКИХ УСТРОЙСТВ С МАТЕРИАЛАМИ С ФАЗОВЫМ ПЕРЕХОДОМ ДЛЯ ОПТИЧЕСКОЙ ПАМЯТИ

Зубкова А.И., Пуговкин М.Д., Смирнов Н.В.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

Работа посвящена численному расчету распространения ТЕ - оптической моды инфракрасного излучения длины волны 1550 нм и поглощения в волноводе из нитрида кремния Si_3N_4 , содержащем тонкую пленку халькогенидного стеклообразного полупроводника $\text{Ge}_2\text{Sb}_2\text{Te}_5$ (GST). Аморфная и кристаллическая фазы GST структурно различаются, что приводит к значительным различиям в оптических и электрических свойствах. Это находит практическое применение в устройствах хранения информации. Особенно заметна разница показателей преломления этих фазовых состояний в телекоммуникационном диапазоне длин волн (1,55 - 1,62 мкм). В работе были подобраны необходимые геометрические параметры структур и проведено моделирование в ПО Comsol с использованием 2D и 3D моделей. Было выявлено, что при увеличении высоты и ширины GST увеличивается поглощение оптического излучения как в аморфном, так и в кристаллических состояниях. Также оптическая мода при аморфном состоянии GST остается в волноводе, в отличие от кристаллического.

Введение

Использование фотонных интегральных схем (PIC) получает большое распространение благодаря увеличению скорости передачи информации и пропускной способности, а также из-за преодоления недостатков традиционных электронных интегральных схем, требующих больших затрат времени и энергии [1,2].

Несмотря на эффективность использования PIC, по-прежнему сталкиваются с рядом технических проблем, таких как большая занимаемая площадь, высокое энергопотребление и нехватка оптической памяти, возникающие из-за методов активной настройки, используемых для управления оптическими излучением. Эти проблемы можно частично решить, объединив халькогенидные материалы с фазовым переходом (PCM), такие как $\text{Ge}_2\text{Sb}_2\text{Te}_5$ (GST), с волноводом из нитрида кремния Si_3N_4 , что особенно применимо в приложениях с реконфигурируемыми оптическими схемами из-за энергонезависимой природы GST. PCM была предложена для реконфигурируемых и энергонезависимых устройств из-за огромного оптического контраста между аморфным и кристаллическим состояниями PCM [3]. Изменение состояния GST приводит к изменению параметров излучения в волноводе.

Для понимания взаимодействия между GST и кремниевым волноводом необходимо провести ана-

лиз спектров поглощения, отражения и пропускания структур, что в свою очередь позволяет определить коэффициент затухания аморфных и кристаллических областей GST от размеров структур.

Целью нашей работы является расчет ТЕ-моды в двумерном сечении планарного волновода для различных размеров области GST. Для этого нами было проведено моделирование в ПО Comsol зависимости спектров пропускания и поглощения длин волн в ближнем ИК-диапазоне, используемым в телекоммуникациях [4].

Физические основы принципа работы

В основе работы оптических волноводов из нитрида кремния лежит принцип полного внутреннего отражения, при котором излучение отражается от границ раздела сред и распространяется в среде с большим показателем преломления и допустимыми потерями мощности излучения (Рис. 1).

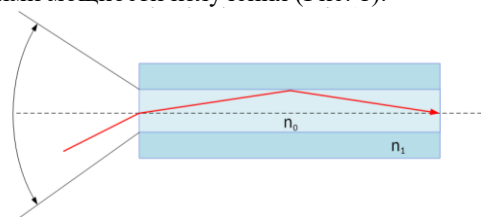


Рис. 1. Распространение света в волноводе

Величина группового показателя преломления n_g :

$$n_g = n_{eff} - \lambda_0 \frac{dn_{eff}}{d\lambda_0} \quad (1.1)$$

где n_{eff} – эффективный показатель преломления:

$$n_{eff} = n_0 \sin(\alpha) \quad (1.2)$$

λ_0 – длина волны,

n_0 – коэффициент преломления оптического волновода,

α – угол падения света на границу раздела сред.

Моделирование

Моделирование производилось в ПО Comsol с использованием 2D и 3D моделей, а также свипирования параметров. Для каждого случая была рассчитана ТЕ-мода в двумерном сечении планарного волновода из нитрида кремния, покрытого слоем аморфного GST и кристаллического GST.

Результаты расчета распространяемого электромагнитного поля в волноводе представлены на рисунках 2 и 3 для кристаллического и аморфного состояния соответственно.

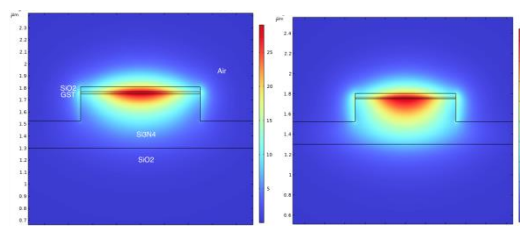


Рис. 2. Распространение моды в волноводе с тонким слоем GST в кристаллическом состоянии

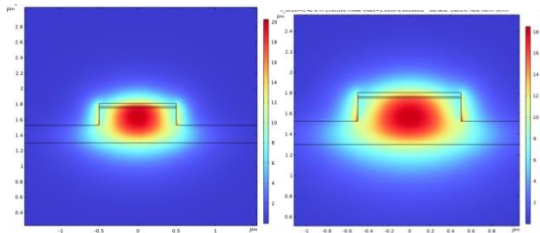


Рис. 3. Распространение моды в волноводе с тонким слоем GST в аморфном состоянии

Нами были выбраны следующие геометрические параметры структур, которые представлены на рис. 4

Name	Expression	Value
h_SiO2	2.6[um]	2.6E-6 m
h_Si3N4	450[nm]	4.5E-7 m
w_Si3N4	1[um]	1E-6 m
h_G	20[nm]	2E-8 m
w_SiO2	7[um]	7E-6 m
H_Si3N4_2	h_Si3N4/2	2.25E-7 m
wl	1550[nm]	1.55E-6 m
n_Si3N4	$\sqrt{1+(2.8939 \cdot (w/1.9793))}$	
n_SiO2	$\sqrt{1+0.6961663 \cdot 1.444}$	
n_GST_AM	n_GST_AM(wl)	4.2282
k_GST_AM	k_GST_AM(wl)	0.046422
n_GST_cryst	n_GST_cryst(wl)	6.467
k_GST_cryst	k_GST_cryst(wl)	0.92037
h_SiO2_cl	40[nm]	4E-8 m
w_G	1000[nm]	1E-6 m

Рис. 4. Геометрические параметры структур

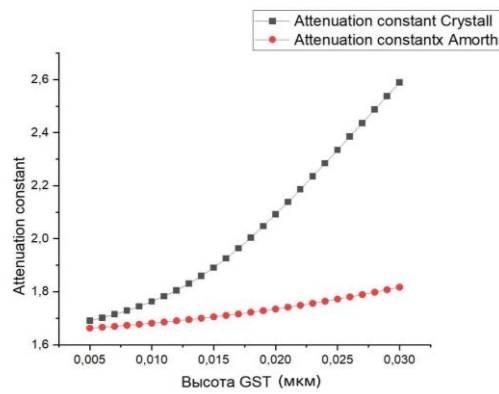


Рис. 5. Зависимость поглощения от высоты GST

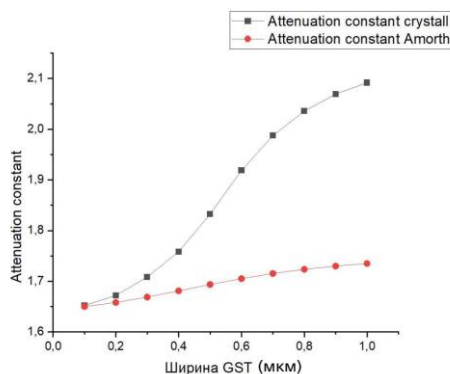


Рис. 6. Зависимость поглощения от ширины GST

Мы получили следующий результат - из-за малого значения действительной части показателя преломления аморфного-GST оптическая мода в основном остается в волноводе, в отличие от волновода, покрытого кристаллическим-GST, который обеспечивает сильный захват TE-моды слоем GST.

На рисунках 5 и 6 можно видеть зависимость поглощения (AttenuationConstant) от ширины и высоты волновода для аморфного и кристаллического состояния GST пленки.

Как следует из проведенных расчетов, при увеличении размеров как ширины, так и высоты GST увеличивается поглощение и в кристаллическом, и в аморфном состоянии, однако при кристаллическом рост поглощения происходит намного сильнее.

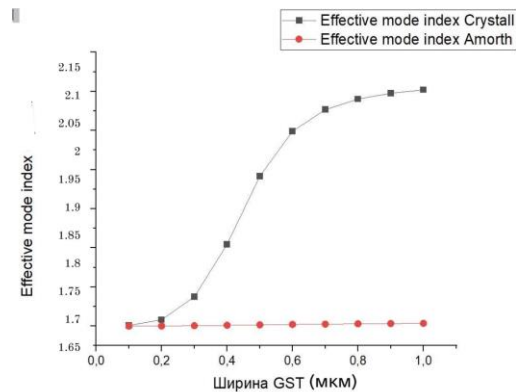


Рис. 7. Зависимость эффективного показателя преломления от ширины GST

Были рассчитаны зависимости эффективного показателя преломления (n_{eff}) от ширины и высоты GST в аморфном и кристаллическом состояниях. Зависимости представлены на рисунках 7 и 8.

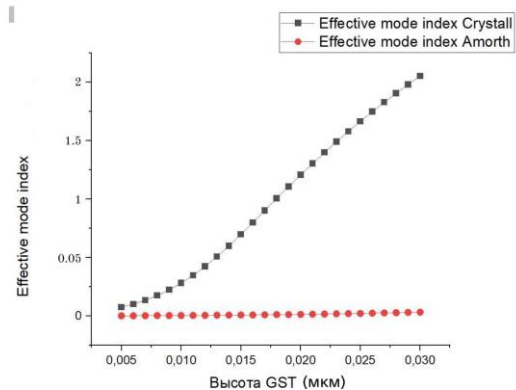


Рис. 8. Зависимость эффективного показателя преломления от высоты GST

Было установлено, что при увеличении ширины и высоты GST увеличивается эффективный показатель преломления в кристаллическом состоянии, а в аморфном состоянии практически не изменяется.

Полученные результаты позволят лучше понять каким образом использовать GST для управления параметрами сигнала волноводных структур. Чем больше поглощение, тем меньшая интенсивность оптического излучения способна нагреть GST до температуры, при которой происходит переход между кристаллической и аморфной фазами при включении прибора.

В настоящее время можно получить тонкие пленки GST с несколькими фазами и различным соотношением аморфного и кристаллического объемов. Чем больше размер слоя GST на волноводе, тем больше оптический эффект переключения и оптический контраст. Однако, уменьшение размера слоя GST потенциально может привести к увеличению энергоэффективности устройства, что открывает путь для создания сотни и тысячи устройств на одном кристалле

По этой причине коэффициент поглощения является одним из ключевых параметров при разработке нанопотонных устройств.

Заключение

С помощью ПО Comsol мы смоделировали распределение моды в волноводе, который содержит материал с фазовым переходом на поверхности. Были получены зависимости поглощения инфракрасного излучения длины волны 1550 нм и эффективного показателя преломления в аморфном и кристаллическом состояниях от ширины и высоты слоя GST. Из результатов моделирования мы можем сделать вывод, что при аморфном состоянии GST оптическая мода остается в волноводе, в отличие от кристаллического состояния GST. Так же при изменении ширины и высоты GST увеличивается поглощение оптического излучения в обоих состояниях GST, однако в кристаллическом состоянии рост поглощения происходит заметно сильнее.

Список литературы

1. Wu C. Programmable phase-change metasurfaces on waveguides for multimode photonic convolutional neural network/ Wu C., Yu H., Lee S., Peng R., Takeuchi I., Li M. - //Nat. Commun. 2021, 12, 96.
2. Feldmann J. Integrated 256 Cell Photonic Phase-Change Memory With 512-Bit Capacity/Feldmann, J., Youngblood N., Li X., Wright, C.D., Bhaskaran, H., Pernice W.H.P. - //IEEE J. Sel. Top. Quantum Electron. 2020, 26, 1–7.
3. Wang J. Overview of Phase-Change Materials Based Photonic Devices/ Wang J., Wang L., Liu J. - //IEEE Access 2020, 8, 121211–121245.
4. Lazarenko P. Size effect of the Ge₂Sb₂Te₅ cell atop the silicon nitride O-ring resonator on the attenuation coefficient/ Lazarenko P., Kovalyuk V., Pavel An P., Prokhodtsov A., Golikov A., Sherchenkov A., Kozyukhin S., Fradkin I., Chulkova G., Gregory Goltsman - //APL Materials 9, 121104 (2021).

РАЗРАБОТКА МАКРОМОДЕЛИ ОПЕРАЦИОННОГО УСИЛИТЕЛЯ С УЧЕТОМ ВЛИЯНИЯ РАДИАЦИИ

Савочкин В.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В данной работе представлены результаты разработки макромодели для ОУ на биполярных транзисторах с учетом влияния дозы ионизирующего излучения от 0 до 3,5 Мрад.

Введение

Операционные усилители широко используются в современных аналого-цифровых схемах. Для схем, используемых в космической и другой специальной аппаратуре, необходимо учитывать влияние факторов накопленной дозы на характеристики схем, в том числе и на ОУ. Макромодели ОУ (рис. 1) широко используются в системах проектирования схем, т.к. они существенно снижают затраты времени моделирования по сравнению с использованием модели ОУ на уровне транзисторов (рис. 2). Как видно, на рис. 1 присутствуют лишь два активных элемента, это транзисторы Q1 и Q2 все остальные элементы являются пассивными, и не требуют больших вычислительных мощностей, в отличие от схемы на рис. 2, где программе необходимо будет посчитать каждый p-n переход на транзисторах и их взаимодействие между собой, что в свою очередь потребует гораздо большее количество времени и вычислительных мощностей. К тому же в отличие от SPICE моделирования на уровне транзисторов, которое требует важных знаний технологических данных по схеме транзисторов, которые, как правило, не предоставляются в открытый доступ или являются коммерческим секретом для моделирования макромоделей не требуется таких данных, достаточно данных приведенных в документации к ОУ, которая имеется в открытом доступе.

Стандартные макромодели ОУ в системах SPICE анализа не учитывают влияние факторов радиации на характеристики ОУ.

Данная работа посвящена разработке SPICE макромодели ОУ на биполярных транзисторах с учетом влияния поглощенной дозы излучения.

Научная новизна данной работы подтверждается отсутствием стандартных макромоделей ОУ с учетом влияния радиации. Кроме того, анализ литературы показал, что для уменьшения вычислительных мощностей в большинстве работ используется поведенческое моделирование ОУ с использованием VHDL-AMS (язык поведенческого описания) моделирования. В работе [1] была разработана поведенческая модель для LM124, моделирующая деградацию параметров ОУ при воздействии ионизирующей радиации от 0 до 190 крад в которой учитывалось

влияние радиации на следующие параметры: входное сопротивление, входное напряжение смещения, входной ток смещения, входной ток смещения, предельная скорость нарастания выходного сигнала, усиление без обратной связи, частотная эффективность усилителя, эффект емкости Миллера, насыщение выходного напряжения, выходное сопротивление. Каждый из этих параметров определяется функцией определенным образом зависящая от величины ионизирующего излучения. Функция была определена исходя из данных полученных при SPICE моделировании на уровне транзисторов. Общий вид функции: $\text{Parameter} = f(\text{pre-rad}) [1 + g(\text{TID})]$. Моделирование схемы триггерного релаксационного осциллятора Шмитта с использованием разработанной поведенческой модели LM124 показывает, что поведенческая модель намного быстрее в сравнении SPICE моделью на уровне транзисторов, при этом достигая хорошей точности моделирования.

В статье [2] проводится аналогичная работа, в ней разработана поведенческая модель ОУ MC14573 учитывающая влияние излучения ионизирующей радиации от 0 до 5 крад. Было смоделировано влияние ионизирующего излучения на следующие параметры ОУ: тока смещения, входного напряжения смещения, коэффициента подавления синфазного сигнала, коэффициента подавления нестабильности питания. Как и в прошлой работе для построения зависимостей параметров использовалась функция. Общий вид функции: $\text{Ptid} = f(\text{dose}, p)$. В этой работе функция была определена исходя из реального опыта облучения ОУ. В указанной работе модель ОУ сравнивалась с реальными испытаниями влияния ионизирующей радиации на ОУ. Как указано в источнике, относительная погрешность для различных параметров не превышает 8%.

В работе [3] было создано три высокоуровневых модели ОУ различных производителей, учитывающие переходные процессы с одиночным событием (ASET), в том числе при двух дозах радиации 20 и 50 крад. Параметрами, зависящими от дозы, являлись ток смещения, скорость нарастания и коэффициент усиления без обратной связи. Эффекты влияния ионизирующего излучения учитываются путем введения изменения электрических параметров от дозы.

В работе [4] приведен пример моделирования ОУ на МОП транзисторах с использованием VHDL-AMS моделирования. В работе было произведено моделирование 4-битного флэш-АЦП с учетом влияния радиации. Скорость моделирования поведенческой модели оказалась 4400 быстрее в сравнении с моделированием на уровне транзисторов с относительной погрешностью ниже 25%.

Также были рассмотрены источники с моделированием влияния других факторов на ОУ. в статье [5] рассматривается моделирование высоких температур на ОУ. Однако, как и в прошлых статьях, там использовалось поведенческое моделирование с использованием VHDL-AMS.

Обзор литературы подтвердил необходимость в создании макромодели ОУ с учетом влияния радиа-

ции. Недостатками перечисленных выше работ является отсутствие значений параметров, зависящих от радиации и отсутствия описания элементов, которыми вводились те или иные параметры деградации ОУ. Кроме того, в приведенных работах используется VHDL-AMS описание, которое совместимо далеко не со всеми SPICE пакетами, в отличие от макромодели, которая будет более совместимой с системами SPICE моделирования.

Целью данной работы является создание макромодели операционных усилителей (ОУ) различного назначения, учитывающих воздействие радиации для проектирования радиационно-стойких перспективных аналоговых и аналого-цифровых схем обработки информации.

Для этого необходимо было решить следующие задачи: провести анализ макромодели ОУ с учетом влияния факторов радиации, исследовать характеристики ОУ при воздействии факторов радиации, отработать методики создания SPICE макромодели ОУ.

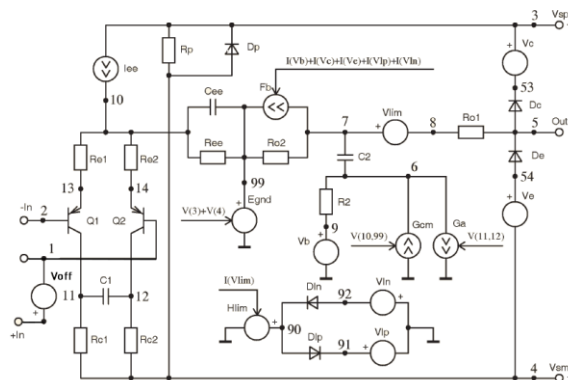


Рис. 1. Вариант схемы макромодели Боyle'a с биполярными p-n-p транзисторами на входе для ОУ

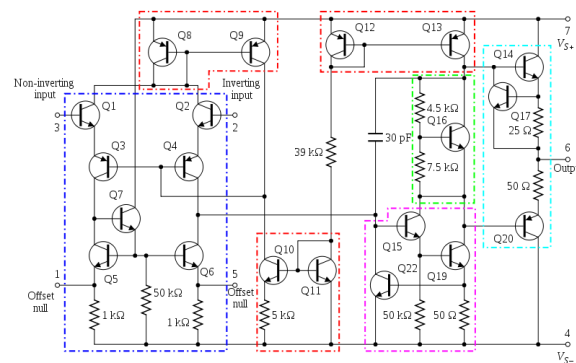


Рис. 2. Интегральная схема операционного усилителя uA741 на биполярных транзисторах

Разработка макромодели ОУ uA741 с учетом влияния дозы облучения

Для отработки методики создания и анализа макромодели ОУ с учетом влияния факторов радиации был выбран операционный усилитель uA741 связи с имеющейся в открытом доступе информацией о влиянии радиационного облучения на характеристики ОУ в работе [6].

Как сказано в работе [6], при облучении данного ОУ поглощенная доза составила $3,5 \times 10^6$ рад. Облучение проводилось с помощью установки для облуче-

ния кобальтом-60 с мощностью дозы 140 рад/сек. Все образцы были облучены при комнатной температуре. В статье [6] было показано, что основными параметрами ОУ, зависящими от дозы, являются: Усиление в дифференциальном режиме (ADM), коэффициент подавления синфазного сигнала (CMRR), Входное сопротивление синфазного сигнала (CMIR), напряжение и ток смещения, граничная частота. Показано, что усиление ОУ является наиболее чувствительной функцией гамма-облучения (рис. 3), где оно теряет более 60 % от своего начального значения после воздействия на ОУ излучения дозой $3,5 \times 10^6$ рад. Как следует из статьи, Коэффициент подавления синфазного сигнала (CMRR) в рассматриваемой работе упал от начального значения в 95 дБ до 0 дБ в результате увеличения дозы гамма-излучения примерно до 2,8 Мрад. Входное сопротивление синфазного сигнала (CMIR) падает с 9 МОм до 1,5 МОм в результате гамма-облучения до 3,5 Мрад. Гамма-облучения вызывают ухудшение свойств входных транзисторах, приводящие к увеличению отличий в их характеристиках, в результате чего входное смещение значения напряжения и тока дифференциального усилителя $\mu A741$ резко увеличивается.

Для данной работы вначале была сделана макро-модель ОУ $\mu A741$ без учета влияния ионизирующего излучения, после чего в данную модель вводились изменения параметров модели для учета влияния ионизирующего излучения.

Для разработки радиационно-зависящей макро-модели ОУ вводились зависимости от дозы следующих элементов схемы рис.1:

- напряжение на источнике V_{off} , включенного на входе макро-модели ОУ, для изменения смещения нуля,
- коэффициент передачи генератора G_a для изменения коэффициента усиления ОУ,
- резистор R_{ee} для изменения подавление синфазного сигнала [7].

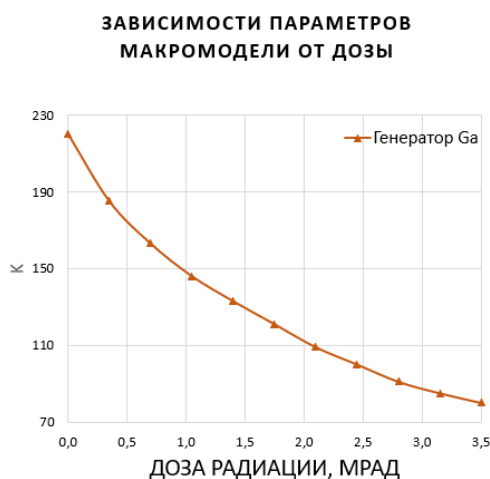


Рис. 3. Полученная зависимость коэффициент передачи генератора G_a для макро-модели

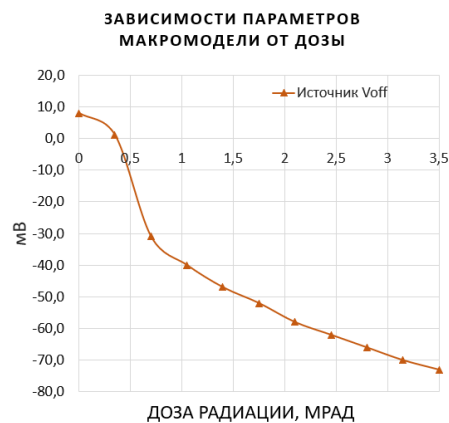


Рис. 4. Полученная зависимость напряжения для источника V_{off}

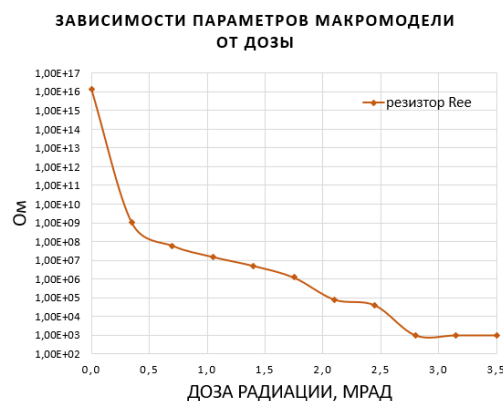


Рис. 5. Полученная зависимость сопротивления резистора R_{ee} для макро-модели

Определение параметров макро-модели производилась путем оптимизации радиационных зависимостей за счет ручной подгонки смоделированных зависимостей под данные, приведенные в статье [6].

На рисунках ниже приведены: графики сравнения результатов моделирования макро-модели с результатами измерений в статье [6] коэффициента усиления (без ОС) и смещение нуля, коэффициент подавления синфазного сигнала, а также полученные зависимости параметров макро-модели.

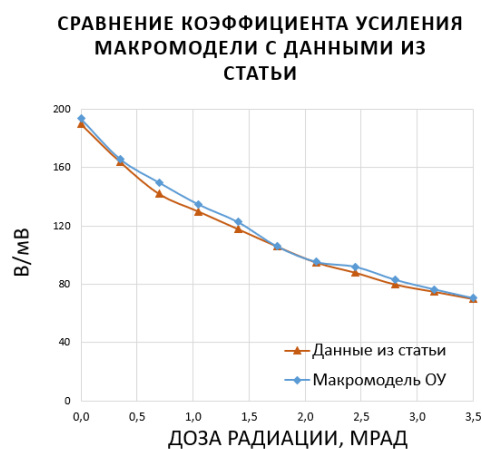


Рис. 6. Сравнение результатов моделирования с помощью макро-модели и измеренной в статье [6] зависимости коэффициента усиления ОУ $\mu A741$ (без ОС) от полученной дозы

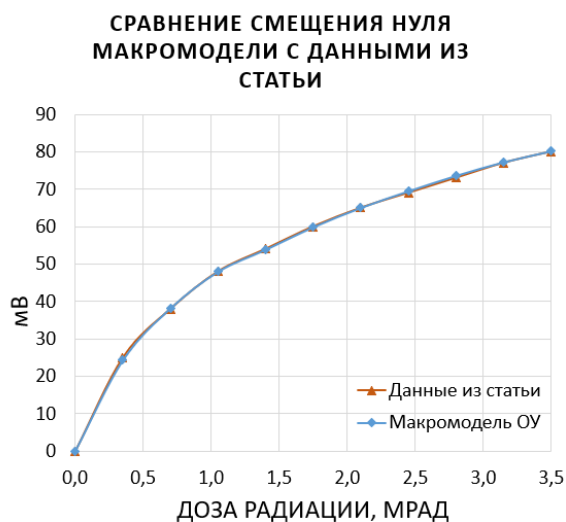


Рис. 7. Сравнение результатов моделирования с помощью макромодели и измеренной в статье [6] зависимости смещения нуля ОУ $\mu A741$ от полученной дозы

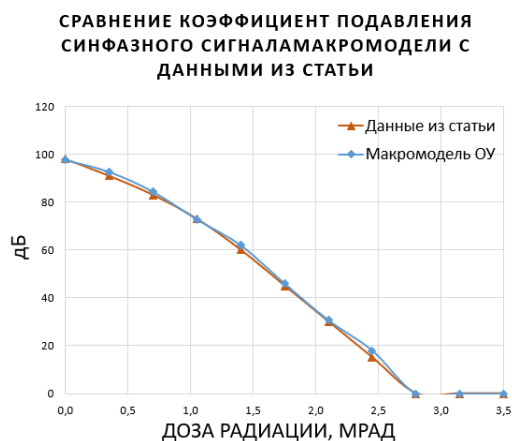


Рис. 8. Сравнение результатов моделирования с помощью макромодели и измеренной в статье [6] зависимости коэффициента подавления синфазного сигнала ОУ $\mu A741$ от полученной дозы

Видно, что результаты моделирования зависимости коэффициента усиления ОУ (без ОС), смещения нуля и коэффициента подавления синфазного сигнала от полученной дозы близки к результатам измерений в работе [6].

Заключение

Была разработана макромодель для ОУ $\mu A741$ на биполярных транзисторах, учитывающая влияние факторов полученной дозы до $3,5 \times 10^6$ рад. Радиационно-зависимыми параметрами макромодели являются: источник V_{off} , резистор R_{ee} , коэффициент передачи генератора G_a . Средняя погрешность моделирования для коэффициента усиления составила 0,03%, для смещения нуля составила 0,63%, для коэффициента подавления синфазного сигнала составило 1,95% по сравнению с результатами измерений, приведенными в литературе [6].

Список литературы

1. Srikanth Jagannathan «Portable Behavioral Modeling of TID Degradation of Voltage Feedback Op-Amps» // Nashville, Tennessee 2009-07-19. -с. 35-54
2. Jin-hui Liu, Quan Wang, Ying Zhang, Gang Liu, Bo Wan « Modeling operation amplifier based on VHDL-AMS for TID effect »// NUCL SCI TECH., 2016. -с. 2-7
3. Fabien Roig, Student Member, IEEE, L.Dusseau, Senior Member, IEEE, A. Khachatryan « Modeling and Investigations on TID-ASET's Synergistic Effect in LM124 Operational Amplifier From Three Different Manufacturers» // IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 60, NO. 6, DECEMBER 2013. -с. 4432-4435
4. Esko Olavi Mikkola; Bert Vermeire; H. G. Parks; Russell Graves « VHDL-AMS Modeling of Total Ionizing Dose Radiation Effects on CMOS Mixed Signal Circuits» // IEEE Transactions on Nuclear Science, August 2007. -с. 930-934
5. Sahbi Baccar, Timothée Levi, Dominique Dallet, Vladimir Shitikov, François Barbara «A Validity Study of an Industrial SPICE-based Opamp Macromodel for High-Temperature Simulation»// IEEE Instrumentation and Measurement Technology Conference May 2012. -с. 3-5
6. F.a.s SOLIMAN, «Operational amplifier type 741: Characterization and radiation effects»// Research Articles January 1, 1993. -с. 15-31
7. G. Boyle, D. Pederson, B. Cohn, J. Solomon. «Macromodeling of integrated circuit operational amplifiers» // IEEE Journal of Solid-state Circuits, 1 December 1974 . -с. 354-362

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТРОЕНИЯ МОДЕЛИ МОП-ТРАНЗИСТОРА

Ларионов С.А., Уланова А.В.

Национальный исследовательский ядерный университет «МИФИ»,
кафедра №3 «Электроника»,
Институт нанотехнологий в электронике,
спинтронике и фотонике

Аннотация

В работе представлено применение методов машинного обучения, таких как градиентный бустинг и искусственные нейронные сети, для моделирования полупроводниковых приборов на примере моделирования вольтамперной характеристики МОП-транзистора. Данные, использованные для обучения и валидации моделей, рассмотренных в работе, получены с помощью SPICE-моделирования. Продемонстрировано, что применение методов машинного обучения является простым, и при этом достаточно точным, способом моделирования МОП-транзистора.

Введение

Компактными моделями называются модели транзисторов, предназначенные для схемотехнического проектирования электрических цепей и интегральных схем с помощью SPICE-подобных программ моделирования [1].

Современные компактные модели МОП-транзисторов основаны на учёте физических принципов работы полупроводниковых транзисторов и имеют более тысячи параметров. Процедура экстракции этих параметров проводится с помощью специализированных САПР, является довольно трудоёмкой и основана на обработке значительного объёма данных, полученных в результате измерения тестовых структур.

Методы машинного обучения хорошо зарекомендовали себя в построении точных моделей, основанных на эмпирических данных. Они могут отражать нелинейные характеристики полупроводниковых приборов, и, таким образом, машинное обучение может быть использовано для построения модели МОП-транзистора. В частности, такие модели уже были построены на основе искусственных нейронных сетей (ИНС) [2, 3].

В настоящей работе на примере моделирования вольтамперной характеристики МОП-транзистора показано, что решения, основанные на применении методов машинного обучения, таких как градиентный бустинг и ИНС, значительно проще с точки зрения экстракции параметров модели.

Генерация набора данных для обучения

Наборы данных для обучения и валидации (датасеты) были сгенерированы с использованием SPICE-модели [4] BSIM3 [5] и симулятора NGSPICE. Независимыми переменными, то есть входными параметрами моделей, являются: напряжение сток-исток (V_{ds}), напряжение затвор-исток (V_{gs}), ширина канала (W), длина канала (L) и температура окружающей среды (T). Целевой переменной, которую должна предсказать модель, является ток сток-исток (I_{ds}).

Датасет должен был имитировать результаты измерения тестовых структур. Так как разброс параметров технологического процесса изготовления полупроводниковых структур и погрешность средств измерения, приводят к отклонению результатов измерений характеристик МОП-транзистора от результатов, полученных с помощью SPICE-модели, то при генерации датасета, использованного для обучения моделей, к данным был добавлен случайный шум. Область обучения представляет собой набор из 9 различных транзисторов (W , L) (см. рис. 1) и набор различных входных напряжений (V_{ds} , V_{gs}) в диапазоне от 0 В до 3 В с шагом 0,25 В.

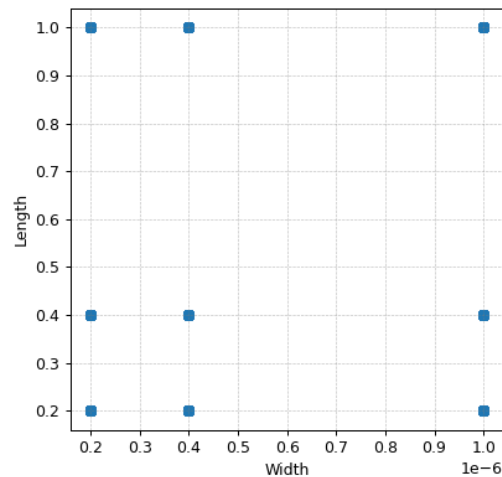


Рис. 1. Распределение точек по длине и ширине канала

Используемые метрики

В настоящей работе для оценивания качества обученных моделей применялись следующие метрики: среднеквадратическая ошибка (MSE), средняя абсолютная ошибка в процентах (MAPE) и коэффициент детерминации (R^2).

MSE является одной из самых популярных метрик в задачах регрессии. MSE применяется в ситуациях, когда необходимо выбрать модель, которая даёт меньше больших ошибок прогноза:

$$MSE(y^{true}, y^{pred}) = \frac{1}{N} \sum_{i=1}^N (y_i^{true} - y_i^{pred})^2,$$

где y_i^{true} – фактический ожидаемый результат и y_i^{pred} – прогноз модели.

MAPE используется для оценки относительной ошибки:

$$MAPE(y^{true}, y^{pred}) = \frac{1}{N} \sum_{i=1}^N \frac{|y_i^{true} - y_i^{pred}|}{|y_i^{true}|}$$

R^2 используется для оценки производительности модели машинного обучения на основе регрессии. Стоит заметить, что у идеальной решающей модели R^2 равен 1.

$$R^2(y^{true}, y^{pred}) = 1 - \frac{1}{N} \sum_{i=1}^N \frac{(y_i^{true} - y_i^{pred})^2}{|y_i^{true} - \bar{y}|^2},$$

где \bar{y} – выборочное среднее.

Реализация моделей градиентного бустинга

Создание моделей МОП-транзистора методами градиентного бустинга выполнялись с использованием открытой программной библиотеки CatBoost [6].

К набору обучающих данных применялась предобработка методом нормализации или методом применения десятичного логарифма.

На рис. 2 представлены регрессионные функции, полученные вне области обучения моделей после различной предобработки.

Оценка погрешности для модели с предобработкой методом нормализации:

- $MSE = 3.290 \times 10^{-9}$;
- $MAPE > 600\%$.

Оценка погрешности для модели с преобразованием десятичным логарифмом:

- $MSE = 9.607 \times 10^{-10}$;
- $MAPE \sim 40\%$.

Из рисунка 2 видно, что при низких напряжениях V_{gs} (рис. 2а) модель, полученная с помощью метода градиентного бустинга, демонстрирует неудовлетворительные результаты, однако в области высоких напряжений V_{gs} (рис. 2б) предсказания модели близки к фактическим значениям, полученным с помощью SPICE-моделирования.

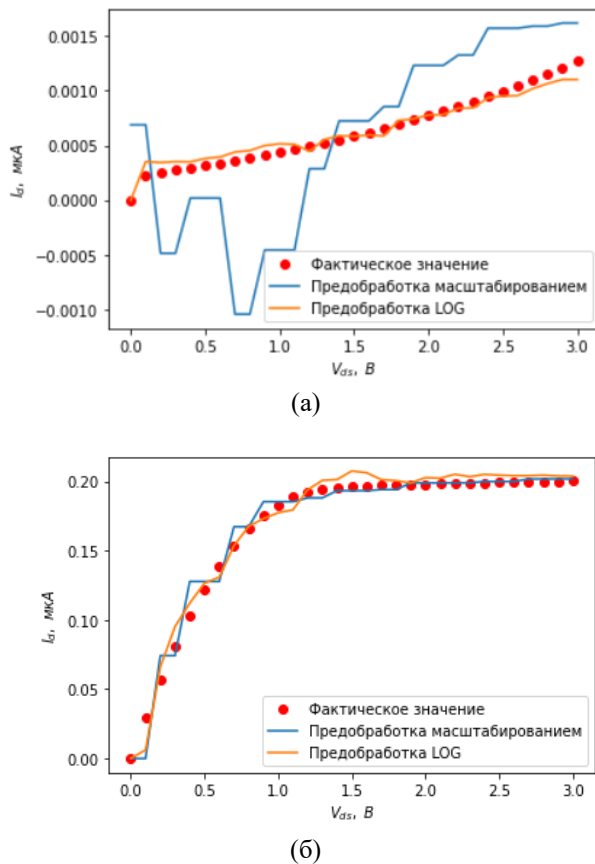


Рис. 2. Примеры работоспособности полученных моделей (для $V_{ds}=0.5$ В и $V_{gs}=2.0$ В)

Реализация модели ИНС

Создание моделей МОП-транзистора методами искусственных нейронных сетей (ИНС) выполнялась с использованием открытой программной библиотеки TensorFlow [7] и Scikit-learn [8].

В качестве входных данных использовались пять различных параметров: V_{ds} , V_{gs} , W , L и T . Входные данные предварительно обрабатывались. Выходной сигнал (I_{ds}) проходил постобработку. Предобработка заключалась в преобразовании на входные данные десятичного логарифма, а затем их нормализации.

Входные данные были разделены на тестовую и обучающую выборку в соотношении 1 к 4 соответственно.

В качестве функции активации в скрытых слоях были использованы функции гиперболического тангенса (Tanh) с применением регуляризации. Размеры скрытых слоёв составляли 20 и 20 нейронов (5-20-20-1). Все слои полностью связаны. В качестве функции потерь использовалось MSE.

Оценка погрешности составила:

- $MSE = 3.969 \times 10^{-9}$;
- $MAPE = 13\%$ (в среднем 16%);
- $R^2 = 0.85$.

Конструирование признаков для реализации модели ИНС

Конструирование признаков (feature engineering) является хорошо известным способом улучшения предсказательной способности моделей машинного обучения. Данный подход заключается в создании новых независимых переменных, являющихся комбинациями уже существующих. Такие новые независимые переменные, к примеру, могут являться суммой или произведением существующих независимых переменных, быть полученными возведением существующей независимой переменной в степень и т.д. [9]

С другой стороны, взаимосвязь между входными и выходными параметрами МОП-транзистора довольно сложна и описывается нелинейными уравнениями. Например, ток I_{ds} в режиме сильной инверсии, который характеризуется тем, что напряжение затвор-исток превышает величину порогового напряжения транзистора на 200 мВ и выше, описывается выражением [10]:

$$I_{ds} = \frac{\mu C_{ox} W}{2 L} (V_{gs} - V_{th})^2,$$

где μ – подвижность основных носителей, C_{ox} – удельная ёмкость подзатворного диэлектрика, V_{th} – пороговое напряжение. Легко заметить, что в выражение для I_{ds} входят V_{gs}^2 и $1/L$.

Таким образом, можно предположить, что качество модели МОП-транзистора на основе ИНС можно улучшить, добавив в модель признаки (независимые переменные), являющиеся комбинациями исходных признаков.

В настоящей работе такая модель МОП-транзистора на основе ИНС была реализована. Дополнительные признаки были сформированы на основе анализа уравнений, описывающих физические принципы работы МОП-транзистора. Данные параметры носят прогностический характер, которые в противном случае ИНС должна извлекать из обучающих данных самостоятельно, например, как в предыдущей реализации модели ИНС.

Архитектура предложенной модели представлена на рис. 3.

Модель ИНС начинается со входного слоя из пяти входных параметров (V_{ds} , V_{gs} , W , L и T), которые

обычно считаются предсказуемыми для выходного параметра I_{ds} .

На этапе генерации к первоначальным входным параметрам, которые прошли предобработку, добавляется набор из 18 признаков с объяснительной силой, полученных разложением общих аналитических уравнений на полиномы V_{ds}^2 , V_{gs}^2 , $V_{ds} V_{gs}$, $V_{ds}^2 V_{gs}$, $V_{ds} V_{gs}^2$, V_{ds}/L , V_{gs}/L , V_{ds}^2/L , V_{gs}^2/L , $V_{ds} V_{gs}/L$, $V_{ds} W$, $V_{gs} W$, $V_{ds}^2 W$, $V_{gs}^2 W$, $V_{ds} V_{gs} W$, $1/T$ и др.

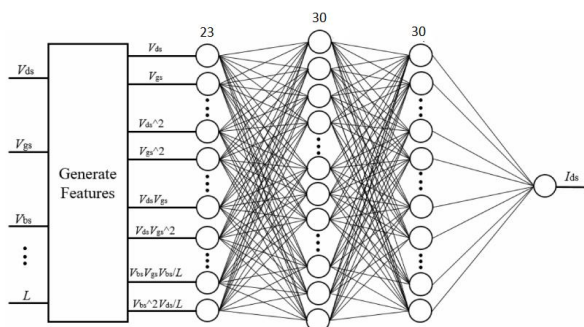


Рис. 3. Архитектура модели ИНС с входными полиномами

ИНС состоит из входного слоя, двух скрытых слоёв и выходного слоя (23-30-30-1). В качестве функции активации в скрытых слоях используются функции гиперболического тангенса (Tanh) с применением регуляризации. Все слои полностью связаны.

Два скрытых слоя обеспечивают гибкость сети для захвата основных взаимосвязей между предоставленными функциями данных, а также для захвата любых дополнительных функций, которые не были предусмотрены на уровне генерации функций. В качестве функции потерь использовалась функция MSE.

Оценка погрешности данной модели ИНС составила:

- $MSE = 3.815 \times 10^{-9}$;
- $MAPE = 10\%$ (в среднем - 11%);
- $R^2 = 0.85$.

Таким образом, предсказательная способность модели ИНС с добавлением полиномов улучшилась по метрикам MSE и MAPE.

Заключение

В настоящей работе было представлено применение методов машинного обучения, таких как градиентный бустинг и искусственные нейронные сети, для моделирования полупроводниковых приборов на примере моделирования вольтамперной характеристики МОП-транзистора. В работе использовались данные, полученные с помощью SPICE-моделирование, имитирующие измерения тестовых структур.

Были реализованы: модель на основе метода градиентного бустинга с различной предобработкой данных, модель на основе ИНС с различной предобработкой данных, а также модель ИНС с генерацией дополнительных входных параметров в виде полиномов.

Модель, реализованная с помощью метода градиентного бустинга, продемонстрировала приемлемые результаты только при высоких значениях напряжений затвор-исток.

Модель на основе ИНС позволяет получить более точный результаты, по сравнению с моделью на основе градиентного бустинга, в соответствии с метриками MSE и MAPE, а конструирование дополнительных признаков в виде полиномов позволяет улучшить предсказательную способность модели данного типа.

Надо отметить, что обучение моделей градиентного бустинга вычислительно эффективнее обучения моделей ИНС.

Можно заключить, что применение методов машинного обучения для моделирования ВАХ МОП-транзистора является довольно простым способом, относительно моделей, основанных на физике.

Список литературы

1. Денисенко В. Принципы построения компактных моделей МОП-транзисторов // Компоненты и технологии. – 2009. – №. 99. – С. 153-158.
2. Hayati M., Rezaei A., Seifi M. CNT-MOSFET modeling based on artificial neural network: Application to simulation of nanoscale circuits // Solid-State Electronics. – 2010. – Т. 54. – №. 1. – С. 52-57.
3. Woo S. M. et al. Machine-Learning-Based Compact Modeling for Sub-3-nm-Node Emerging Transistors // Electronics. – 2022. – Т. 11. – №. 17. – С. 2761.
4. Predictive Technology Model (PTM) // URL: <https://ptm.asu.edu/>.
5. Berkeley Short-channel IGFET Model 3 // URL: <http://bsim.berkeley.edu/models/bsim3/>
6. CatBoost is a high-performance open source library for gradient boosting // URL: <https://catboost.ai/>.
7. TensorFlow is an open source platform for machine learning. // URL: <https://www.tensorflow.org/>
8. Scikit-learn is a Python module for machine learning // URL: <https://scikit-learn.org/>
9. Habal H., Tsonev D., Schweikardt M. Compact models for initial MOSFET sizing based on higher-order artificial neural networks // 2020 ACM/IEEE 2nd Workshop on Machine Learning for CAD (MLCAD). – IEEE, 2020. – С. 111-116.
10. Эннс В. И. Проектирование аналоговых КМОП-микросхем: краткий справочник разработчика / В.И. Эннс, Ю.М. Кобзев; под ред. В.И. Эннса. – М.: Горячая линия-Телеком, 2005.

ВРЕМЯ РЕЛАКСАЦИИ ЭНЕРГИИ В ЭПИТАКСИАЛЬНЫХ TiN ПЛЕНКАХ

*Корчагинская С. А., Баева Э.М.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии,
базовая кафедра квантовой оптики и
телекоммуникаций ЗАО "Сконтел"
МИЭМ НИУ ВШЭ*

Аннотация

В работе измеряются вольт-амперные характеристики сверхпроводниковой TiN пленки в магнитном поле, чтобы получить информацию о времени релаксации энергии. Сравнивая предварительные результаты с результатами полученными другими методами, наблюдается различие в значениях на порядок. Полученные результаты указывают на необходимость аккуратной проверки различных подходов по извлечению времени релаксации.

Введение

Неупругое рассеяние электронов в нормальном состоянии сверхпроводников играет существенную роль в неравновесной физике сверхпроводящих детекторов. Скорости процессов электрон-электронного и электрон-фононного рассеяния вместе контролируют термализацию и охлаждение неравновесных квазичастиц и, следовательно, определяют временные показатели (т.е. джиттер, мертвое время) сверхпроводящих однофотонных детекторов, а также их способность обнаруживать одиночные фотоны [1]. Также эти процессы определяют полосу промежуточных частот болометров на горячих электронах [2] и масштабы скоростей рекомбинации квазичастиц в микроволновых детекторах кинетической индуктивности [3].

Время неупругой релаксации τ_{in} в сверхпроводниковых пленках изучается двумя способами. Первый – через изучение усредненных по времени транспортных характеристик (магнитосопротивление и сопротивление [4], напряжение [5], шум напряжения на низкой частоте [6]). Несмотря на то, что данный подход зарекомендовал себя с 80-х годов, как основной инструмент для исследования неупругой релаксации в материалах, основной его недостаток заключается в аккуратности применений теорий к усредненным по времени характеристикам для определения временной характеристики τ_{in} . Второй подход – изучение спектральных характеристик сопротивления или шума напряжения на сверхпроводящем переходе и теплосопровитлений в нормальном состоянии и др. [7-9]. Этот подход имеет один явный недостаток, который заключается в том, что при интерпретации спектральных характеристик, полученных на переходе сверхпроводник-нормальный металл, рассмотрение системы оказывается сильно упрощено [7]. Для Nb и NbN пленок [10-11] было показано, что оба подхода согласуются между собой. Однако для некоторых пленок не наблюдалось со-

гласия между двумя подходами [5,8-9]. Например, в тонких TiN пленках результаты по времени τ_{in} отличаются на порядок [5,8-9]. Данное разногласие может быть связано с тем, что изученные пленки в работах [5,8-9] отличались по своим свойствам.

Для проверки различных подходов по извлечению времени релаксации, мы дополняем данные по τ_{in} , полученные для эпитаксиальных TiN пленок с помощью шумовой спектроскопии на сверхпроводящем переходе [8]. В данной работе мы исследуем τ_{in} в эпитаксиальных TiN пленках, используя теорию Ларкина-Овчинникова (Л-О) [12]. Измеряя вольт-амперные характеристики в магнитном поле, мы достаем максимальную скорость потока вихрей. Из скорости потока вихрей мы оцениваем τ_{in} в рамках оригинальной Л-О теории [12] и сравниваем с результатами второго подхода.

Экспериментальное получение вольт-амперных характеристик

Образец изготовлен из эпитаксиальной пленки TiN толщиной 9 нм с низким уровнем беспорядка ($k_F l \sim 350$) [13]. Из пленки был сформирован четырех-контактный мостик методами плазмохимического травления, оптической литографии и сканирующей электронно-лучевой литографии (см. Рисунок 1). Критическая температура образца и сопротивление на квадрат пленки составляют $T_c = 5$ К и $R_s = 9$ Ом/кв при 6 К. Длина образца L составляет 40 мкм, ширина w составляет 10 мкм.

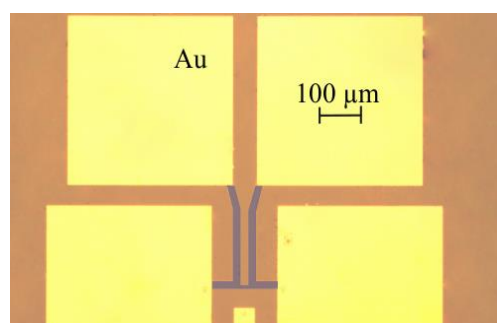


Рис. 1. Фотография образца, полученная оптическим микроскопом

Образец помещается в криостат при температуре 4.2 К. Измерения выполняются в ^4He вставке, внутри которой образец находится в вакууме, снаружи вставки располагается сверхпроводящий соленоид для создания внешнего магнитного поля B ориентированного перпендикулярно образцу. Образец смещается током с помощью источника постоянного тока Yokogawa GS200, напряжение с образца усиливается предусилителем по напряжению Signal recovery 5113. и считывается вольтметром Keysight 2182A/E. На рис. 2 представлена типичная вольт-амперная характеристика образца в конечном магнитном поле B . Видно, что при увеличении тока напряжение монотонно возрастает, а при достижении критического тока, наблюдается скачок напряжения на линейную зависимость. Прирост напряже-

ния перед критическим током может быть объяснен движением потока вихрей и исследован в рамках Л-О теории, описывающей неустойчивость вольт-амперной характеристики.

Исследование вольт-амперных характеристик в рамках теории Ларкина-Овчинникова

В рамках оригинальной теории Л-О при приложении напряжения V к сверхпроводнику второго рода при температуре близкой к T_c меняется функция распределения квазичастиц внутри ядра вихря и скорость движения вихря v . Вследствие этого, квазичастицы в ядре вихря приобретают энергию, отличную от энергии в равновесном состоянии, и начинают диффундировать из ядра. Диффузия квазичастиц приводит к сжатию вихря, и как следствие, коэффициент вязкости вихревой жидкости, связанный с функцией распределения и щелью сверхпроводника, уменьшается с ростом скорости вихря v . При достижении напряжения $V = V^*$, оно достигается при критическом токе, будет достигнута такая скорость вихрей $v = v^*$, при которой сила вязкого трения будет максимальной. Вольт-амперная характеристика будет иметь неустойчивость, т.е. прямой или N-образный скачок. Из рис. 2 мы определяем V^* как напряжение, при котором происходит скачок на IV-характеристике.

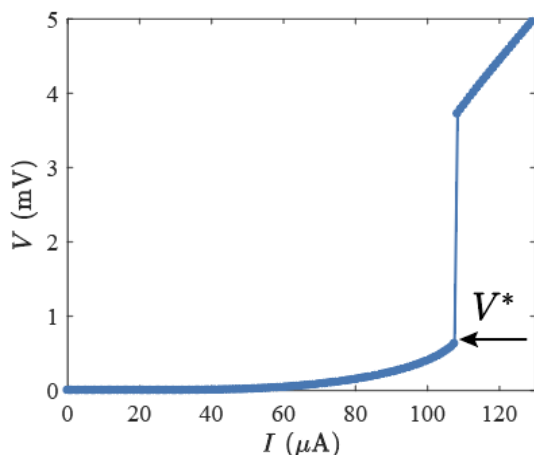


Рис. 2. Вольт-амперная характеристика образца в конечном магнитном поле $B=36$ мТ, измеренная при температуре $T=4.2$ К

Величину V^* мы в дальнейшем строим на рис. 3 в зависимости от приложенного магнитного поля. В малых магнитных полях это напряжение линейно зависит от величины приложенного магнитного поля, что указывает на постоянство скорости v^* движения вихрей. В условиях отсутствия пиннинга вихрей для сверхпроводника в грязном пределе обработка экспериментальных данных осуществляется по следующей формуле [12]:

$$V^* = LB \sqrt{\frac{D}{\pi \tau_{in}}} [14\zeta(3)]^{1/4} \left(1 - \frac{T}{T_c}\right)^{1/4} \quad (1)$$

позволяет нам получить оценку времени релаксации квазичастиц в образце. Коэффициент диффузии D в данной пленке был рассчитан в предыдущих работах из измерений длины когерентности $\xi=22$ нм [8], которая оказывается в два раза ниже, чем в пределе толстой TiN пленки. Данный экспериментальный факт показывает, что тонкие сверхпроводящие TiN пленки находятся в грязном пределе из-за меньшей длины свободного пробега в них [13].

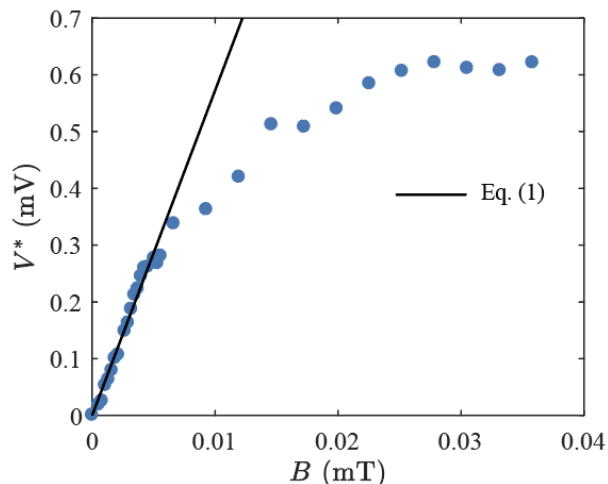


Рис. 3. Зависимость напряжения, при котором происходит скачок вольт-амперной характеристики, от величины внешнего магнитного поля приложенного к образцу

Полученная оценка $\tau_{in} \sim 210$ пс из ур. (1) при $T=0.84T_c$ оказывается в два раза ниже, чем значение, полученное в более разупорядоченных TiN пленках в рамках модифицированной Л-О теории [5], в которой изучалась нелинейная зависимость V^* от магнитного поля. Однако, несмотря на то, что мы пользовались оригинальной Л-О теорией, наши результаты оказались на порядок ниже, чем значения τ_{in} , полученные вторым подходом [8]. Данное отличие может свидетельствовать о неучтенном джоулевым разогреве в низкотокковой части вольт-амперной характеристики из-за использования источника постоянного тока вместо импульсного [14]. Наконец, необходимо в дальнейшем учесть как влияние центров пиннинга на скорость вихрей, так и неравновесную сверхпроводимость [15]. Касательно последней ожидается, что при температуре $T=0.84T_c$ влияние неравновесной заселенности электроподобной и дырочноподобной ветвей энергетического спектра в сверхпроводнике может повлиять на измерение τ_{in} в пределах 10 %.

Исследование выполнено при поддержке гранта Российского научного фонда (проект № 21-79-20084)

Заключение

Предварительные результаты данной работы демонстрируют значительное расхождение двух экспериментальных подходов. Мы надеемся, что в дальнейшем эта неопределенность может быть снята пу-

тем замены источника постоянного тока на импульсный источник тока и проверки влияния пиннинга вихрей на их скорость.

Список литературы

1. Gol'Tsman G.N. et al. Picosecond superconducting single-photon optical detector //Applied physics letters. - 2001. - Vol. 79. - No. 6. - p. 705-707.
2. Semenov A.D. et al. Design and performance of the lattice-cooled hot-electron terahertz mixer //Journal of Applied Physics. - 2000. - Vol. 88. - No. 11. - p. 6758-6767.
3. Leduc H.G. et al. Titanium nitride films for ultra-sensitive microresonator detectors //Applied Physics Letters. - 2010. - Vol. 97. - No. 10. - p. 102509.
4. Lin J.J., Bird J.P. Recent experimental studies of electron dephasing in metal and semiconductor mesoscopic structures //Journal of Physics: Condensed Matter. - 2002. - Vol. 14. - No. 18. - p. R501.
5. Lefloch F., Hoffmann C., Demolliens O. Nonlinear flux flow in TiN superconducting thin film //Physica C: Superconductivity. - 1999. - Vol. 319. - No. 3-4. - p. 258-266.
6. Wellstood F.C., Urbina C., Clarke J. Hot-electron effects in metals //Physical Review B. - 1994. - Vol. 49. - No. 9. - p. 5942.
7. Klapwijk T.M., Semenov A.V. Engineering physics of superconducting hot-electron bolometer mixers //IEEE Transactions on Terahertz Science and Technology. - 2017. - Vol. 7. - No. 6. - p. 627-648.
8. Baeva E.M. et al. Resistance fluctuation spectroscopy of the superconducting transition in epitaxial TiN films //arXiv preprint arXiv:2202.06310. - 2022.
9. Kardakova A. et al. The electron-phonon relaxation time in thin superconducting titanium nitride films //Applied Physics Letters. - 2013. - Vol. 103. - No. 25. - p. 252602.
10. Gershenzon E.M. et al. Electron-phonon interaction in ultrathin Nb films //Sov. Phys. JETP. - 1990. - Vol. 70. - No. 3. - p. 505-511.
11. Lomakin A.I. et al. Evidence of the disorder-independent electron-phonon scattering time in thin NbN films //arXiv preprint arXiv:2207.05012. - 2022.
12. Larkin A.I., Ovchinnikov Y.U. Nonlinear conductivity of superconductors in the mixed state //Sov. Phys. JETP. - 1975. - Vol. 41. - No. 5. - p. 960-965.
13. Saveskul N.A. et al. Superconductivity Behavior in Epitaxial Ti N Films Points to Surface Magnetic Disorder //Physical Review Applied. - 2019. - Vol. 12. - No. 5. - p. 054001.
14. Grimaldi G. et al. Evidence for low-field crossover in the vortex critical velocity of type-II superconducting thin films //Physical Review B. - 2010. - Vol. 82. - No. 2. - p. 024512.
15. Ивлев Б.И. и Копнин Н.Б. Теория токовых состояний в узких сверхпроводящих каналах //УФН. - 1984. - Т. 142. - В. 3. - с. 435.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ЭФФЕКТА ПОЛЯ НА ЭЛЕКТРИЧЕСКИЕ ХАРАКТЕРИСТИКИ СТРУКТУР С УНТ НА ОСНОВЕ ДИОДА ШОТТКИ

Вовк Н.А.^{1,2}, Шибалов М.В.², Мумляков А.М.², Красильников М.И.²,

Тевяшов А.А.², Смирнов К.В.¹, Ануфриев Ю.В.²

¹Национальный исследовательский университет

«Высшая школа экономики»,

департамент электронной инженерии

МИЭМ НИУ ВШЭ,

²Институт нанотехнологий микроэлектроники
Российской академии наук

Аннотация

Исследовано влияние внешнего приложенного электрического поля на характеристики структуры диода Шоттки с массивом горизонтально ориентированных углеродных нанотрубок (УНТ) в качестве функционального слоя. Слой углеродных нанотрубок формировался методом спреевого нанесения с последующей гидрообработкой. Полученные результаты демонстрируют, что одноканальные структуры с УНТ на основе диода Шоттки с металлическими контактами из Ni и Sc имеют ярко выраженную нелинейность при температуре 3К, которая может быть использована в элементах криоэлектроники различного функционального назначения.

Введение

Интегрирование низкоразмерных материалов УНТ в устройства имеет ряд преимуществ, превосходящих кремниевые структуры, такие как высокая подвижность носителей, высокая токопроводящая способность, создание прозрачных слоев, совместимых с гибкими подложками[1]. Полупроводниковая структура УНТ позволяет создавать устройства на основе барьера Шоттки за счет разницы работы выхода между металлом и полупроводником. На основе выпрямляющего контакта металл-полупроводник с использованием одиночных УНТ [2] или массива УНТ [3], уже изготовлен ряд устройств с асимметричными контактами и электродными металлами [4]. Принцип действия работы полупроводниковых полевых транзисторов [5] основан на эффекте поля-управление электрофизическими параметрами поверхности твердого тела с помощью электрического поля, приложенного по нормали к поверхности, с возможностью обогащения или обеднения приповерхностного слоя носителями заряда полупроводника. Введение затворного металла в структуру с массивом УНТ позволяет реализовать эффект поля в устройстве с управлением сопротивления канала, что необходимо для реализации логических интегральных схем и аналоговых [6].

В данной работе продемонстрирована технология создания трехэлектродного устройства с функциональным слоем из УНТ. Исследовалось влияние затворного напряжения на электрические характеристики контакта Шоттки Ni-УНТ-Sc как при комнатной температуре, так и при криогенных.

Методика эксперимента

В качестве подложек использовались кремниевые пластины с термическим оксидом диаметром 4". Образцы были созданы совокупностью технологических методов, включающих в себя осаждение, процесс литографии и травления. В качестве затворного диэлектрика была выбрана пара $\text{HfO}_2/\text{Al}_2\text{O}_3$ толщиной 40 нм для улучшения пробойных характеристик и уменьшения токов утечки на сток Ni и исток Sc (рис.1). Функциональный слой из УНТ формировался методом спреевого нанесения суспензией 100 мл, с последующей обработкой гидромеханическим воздействием поверхности подложки, что позволяет получить однородные слои УНТ с заданным поверхностным сопротивлением [7]. Выбор подходящих электродов является предварительным условием для создания контакта Шоттки. Из литературных источников известно, что работа выхода УНТ варьируется $\phi_{\text{cnt}} = 4.6-5$ эВ [8] и измерена от уровня Ферми полупроводника p-типа. Металл для контакта Шоттки выбирается с большей работой выхода, по сравнению с полупроводником p-типа, для омического контакта с меньшей работой выхода. Исходя из зонной теории, для контакта Шоттки был выбран металл Ni толщиной 100 нм с работой выхода $\phi_{\text{Ni}} = 5.35$ эВ и для омического контакта Sc толщиной 100 нм с работой выхода $\phi_{\text{Sc}} = 2.7$ эВ. Для предотвращения окисления и для улучшения омического контакта на Ni и Sc электроды осаждался слой Al толщиной 100 нм.

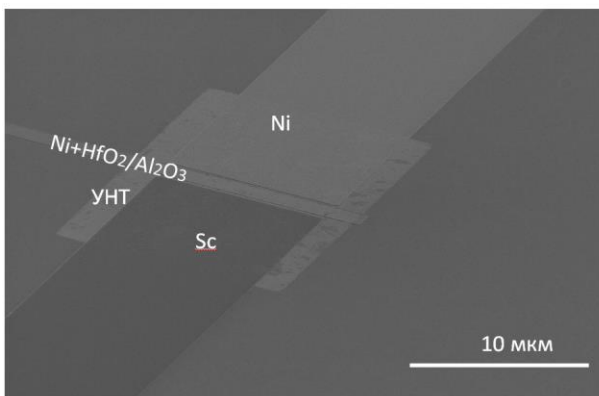


Рис. 1. Снимок сканирующей электронной микроскопии (СЭМ) этапа формирования контактов из Ni и Sc к функциональному слою из УНТ

Методы исследования

Для проведения исследований были изготовлены образцы с различной длиной и шириной канала. Используя малощумящий, прецизионный источник - измеритель 2450 Keithley были получены ВАХ (вольт-амперные характеристики) образцов, которые демонстрируют наилучший эффект поля. Ярко выраженный эффект поля наблюдался на образцах с наибольшей длиной равной 56 мкм и шириной 50 мкм канала (рис.2). Семейство ВАХ, измеренное при различных напряжении на затворе указывает на то, что положительный потенциал на затворе закрывает канал, значит проводящий канал состоит из УНТ с p-

типом проводимости (рис. 2). ВАХ имеют почти симметричную омическую характеристику, что может быть связано с наличием в массиве УНТ многостенных металлических трубок, что вносит существенный вклад в проводимость канала.

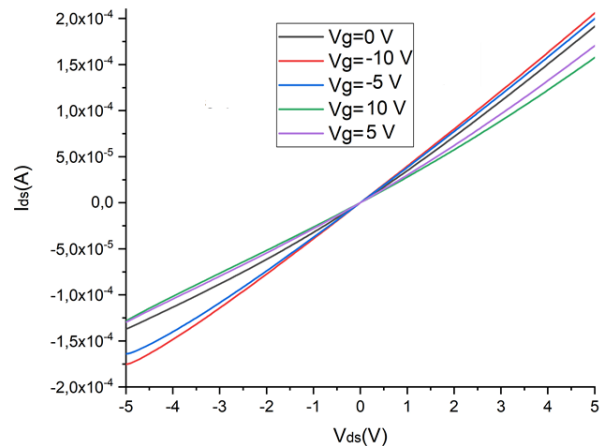


Рис.2. Семейство ВАХ образца в зависимости от приложенного напряжения на затвор

Далее проводились температурные измерения сопротивления образца при фиксированном токе смещения 1 мкА с комнатной температуры до температуры 3К (рис. 3). Измеренная характеристика $R(T)$ типична для УНТ и имеет 3 основных области [9]. Область 1 описывает омическое поведение и слабо зависит от температуры. Область 2, начиная с температуры 100 К и ниже, свидетельствует о полупроводниковом типе проводимости. С понижением температуры концентрация зарядов (электронов и дырок) уменьшается, что приводит к уменьшению числа носителей заряда, доступных для рекомбинации, и, статистически, проводимость УНТ уменьшается, тогда как удельное сопротивление УНТ увеличивается. Авторы источника [9], предполагают, что на данном участке зависимости $R(T)$ механизм барьерной туннельной проводимости доминирует над металлической проводимостью через перколяционные переходы, то есть через проводящие и непроводящие пути. Область 3 может характеризовать вымораживание всех дефектов, как в функциональном слое, так и на контактах.

Сравнение двух ВАХ характеристик при разных температурах демонстрирует сильную нелинейность при температуре 3 К (рис. 4). Помимо неидеальных контактов между металлом истока, стока и УНТ, барьер Шоттки и эффект кулоновской блокады могут рассматриваться как причинами не омического поведения при низких температурах [10]. Также, из-за наличия высоких токов, имеет место Джоулево тепло, когда образец может работать в термисторном режиме при переходе $R(T)$ из области 2 в область 1.

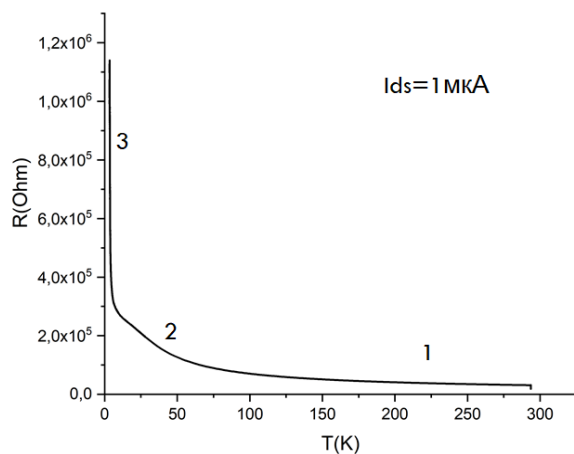


Рис. 3. Зависимость сопротивления образца от температуры при постоянном токе смещения образца 1 мкА

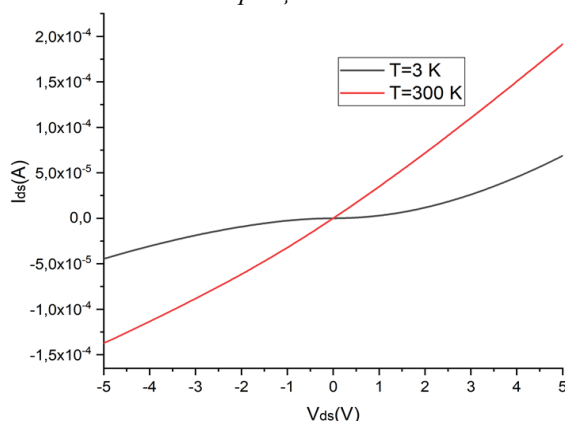


Рис. 4. ВАХ образца измеренные при комнатной температуре (красная кривая) и при рабочей температуре криостата (черная кривая)

Заключение

В данной работе мы успешно изготовили устройства, основанные на контакте Шоттки с встроенным затвором. Были измерены электрические характеристики образцов, демонстрирующие нелинейную зависимость ток-напряжение при низких температурах с влиянием напряжения на затворе. В настоящее время предлагаются подходы к дальнейшему улучшению топологии устройства и его электрических характеристик, что делает возможным изготовление на кристалле структуры с более сильным полевым эффектом в массиве УНТ для фундаментальных исследований, а также для применения в криоэлектронике. Важно отметить возможность использования исследуемых структур при криогенных температурах в связке с сверхпроводящими элементами в качестве СВЧ транзистора на полевым эффекте.

Автор выражает благодарность М.А. Тархову, Н.В. Порохову, Е.Р. Тимофеевой, А.М. Тагаченкову, И.В. Трофимову, Е.В. Зеновой и В.В. Сень за помощь в постановке задачи и полезные обсуждения.

Список литературы

1. Yongwoo L., Haesun J., Bongsik C. et al. // The Royal Society of Chemistry. 2019. V. 9. № 38. P. 22124-22128. DOI: 10.1039/C9RA02855B

2. Yang M. H., Teo K. B. K. and Milne W. I. // Applied Physics Letters. 2005. V. 87. № 25. DOI: 10.1063/1.2149991

3. Byeongju K., Joohyung L., Seon N. et al. // Elsevier B.V. All rights reserved. 2012. V. 169. P. 182-187. DOI: 10.1016/j.snb.2012.04.063

4. Svensson J. and Eleanor E. B. // Journal of applied physics. 2011. V. 110. № 11. P. 111101-16. DOI: 10.1063/1.3664139

5. Hills G., Lau C., Wright A. et al. // Nature. 2019. V.572. № 7771. P.595. DOI: 10.1038/s41586-019-1493-8

6. Hartmann M., Hermann S., PHIL F. et al. // IEEE Journal of microwaves. 2021. V.1. № 1. P.275-287. DOI: 10.1109/JMW.2020.3033781

7. Вовк Н.А., Тархов М.А., Порохов Н.В. et al. // Российские нанотехнологии. 2022. V.17. № 6. P.730-736. DOI: 10.56304/S1992722322060188

8. Liu. P, Sun Q., Zhu F. et al. // Nano Letters. 2008. V.8. № 2. P.647-651. DOI: 10.1021/nl0730817

9. Monea B.F., Ionete E.I, Ionut Spiridon S.I. et al. // Sensors. 2019. V.19. № 11. P.2464-2489. DOI: 10.3390/s19112464

ПОЛУЧЕНИЕ НАНОПРОВОЛОК ИЗ ЖЕЛЕЗА, ИССЛЕДОВАНИЕ ИХ СТРУКТУРЫ

Калачикова И.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

В работе методом матричного синтеза были получены массивы железных нанопроволок (НП) различных диаметров. Исследован характер заполнения матриц, а также структурные особенности получаемых НП. Проведен рентгеноструктурный анализ, который позволил определить тип решетки (ОЦК) и показал закономерное изменение параметра решетки.

Введение

В последние годы особое внимание к себе привлекли нанопроволоки, которые являются одномерными нанообъектами. Массивы НП уже начинают применяться в устройствах магнитной записи информации, в качестве эмиттеров электронов и ионов, в усилении оптических сигналов, в качестве сенсоров. Весьма перспективными является исследования по адресной доставке лекарств с использованием НП

Особый интерес представляют НП из железа. Предполагается, что, варьируя параметры получения железных НП можно изменять их структуру, которая определяет их основные эксплуатационные свойства.

Одним из методов получения НП является матричный синтез. Этот метод основан на заполнении пор заранее подготовленной матрицы выбранным материалом. В настоящей работе в качестве матриц применялись полимерные трековые мембраны (ТМ),

а способом заполнения было выбрано электрохимическое осаждение. Данный метод позволяет широко варьировать структурные, геометрические и соответственно физические свойства получаемых объектов. Однако для получения предсказуемых физических свойств массивов необходимо установить связь режимов получения и геометрических параметров матриц с структурными особенностями получаемых НП.

Целью представленной работы является изучение характера заполнения пор матриц с разными диаметрами пор, а также исследование их структуры методом рентгеновской дифракции.

Эксперимент и результаты

В настоящем исследовании в качестве матрицы используется полимерные трековые мембраны толщиной 12 мкм, с диаметрами пор 30, 65, 100, 200, 400, 600 нм с плотность 10^8 - 10^9 пор на см^2 производства ОИЯИ, г. Дубна.

Перед электрохимическим осаждением на одну из сторон ТМ наносился слой меди. Это необходимо для придания электропроводности матрице, а также закупориванию пор, с одной стороны. В дальнейшем данный слой будет являться катодом в процессе электроосаждения.

Процесс нанесения медного слоя делится на два этапа. Вначале методом вакуумного напыления наносится тонкий медный слой толщиной порядка 50 нм. Далее на напыленный слой осуществляется дополнительное осаждение гальваническим способом слоя меди толщиной 4 мкм.

Далее проводилось электролитическое осаждение железных НП, которое проводили в двухэлектродном режиме. Использовался электролит следующего состава (таблица 1).

Таблица 1. Раствор для осаждения Fe

Элемент	Концентрация, г/л
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	120
H_3BO_3	45
$\text{C}_6\text{H}_8\text{O}_6$	1
$\text{C}_{12}\text{H}_{25}\text{SO}_4\text{Na}$	менее грамма на литр
pH= 2,3	

В данном электролите $\text{C}_6\text{H}_8\text{O}_6$ (аскорбиновая кислота) необходима для предотвращения перехода Fe^{+2} в трехвалентное состояние, а $\text{C}_{12}\text{H}_{25}\text{SO}_4\text{Na}$ (лаурилсульфат натрия) используется для улучшения смачиваемости пор.

В качестве источника тока применялся потенциостат-гальваностат Elins P-2X, в качестве анода использовалось железо. Процесс осаждения проводился в потенциостатическом режиме при потенциале 1 В. Для того, чтобы выявить оптимальное время роста нанопроволок, делается тестовый образец, для которого строится хроноамперограмма полного процесса электрохимического осаждения, при помощи которой определяются границы для времени роста проволоки и пленки железа. Хроноамперограммы для всех образцов приведены на рисунке 1.

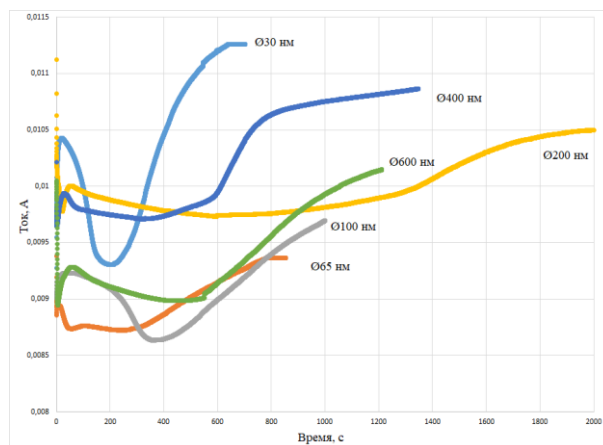


Рис. 1. Хроноамперограммы осаждения железа с диаметрами пор 30, 65, 100, 200, 400 и 600 нм при потенциале 1В

Процесс электроосаждения в поры ТМ можно разделить на два этапа. На первом происходит заполнение поры по длине с образованием структуры в виде нанопроволоки или нити. Далее происходит рост пленки металла на поверхности мембраны. Последней стадии необходимо избегать, поэтому процесс осаждения необходимо контролировать для этого используют две характеристики: время осаждения или заряд протекший через ячейку. [2]

По полученным графикам (рис.1.), для каждого образца было определено время роста НП без выхода осаждаемого металла на поверхность матрицы. Время заполнения представлено в таблице 2.

Таблица 2. Время роста НП

\varnothing , нм	Время роста НП, с
30	200
65	280
100	360
200	700
400	300
600	550

Изучение НП методом электронной микроскопии

При подготовке образцов для их исследования методом СЭМ на установке JEOL JSM-6000PLUS, проводится удаление полимерной матрицы и медного слоя. Для удаления медной подложки применяется раствор на основе аммиака. При удалении трековой мембраны используется концентрированный раствор щелочи NaOH.

На рис. 2 представлен пример получаемых микротографий НП диаметром 200 нм, по нему можно определить, что длина НП совпадает с толщиной полимерной матрицы, а диаметр НП совпадает с диаметром пор в ростовой матрице. Рентгеноспектральный анализ показал, что НП состоят из чистого железа.

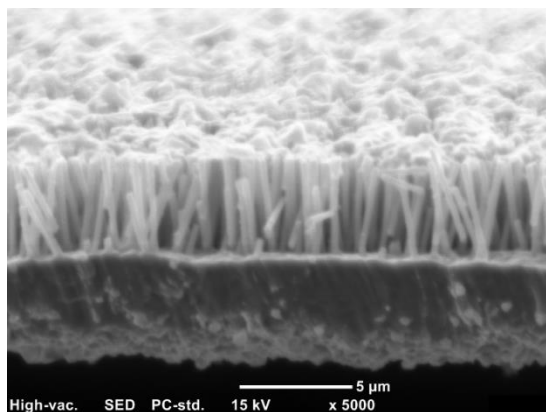


Рис. 2. СЭМ изображение НП диаметром 200 нм

Изучение НП методом рентгеноструктурного анализа

В рамках представленной работы проводились исследования зависимости структуры полученных НП от диаметра пор матрицы методами рентгеноструктурного анализа на установке Rigaku Miniflex 600. (Для этого полимерная матрица и медная подложка не удалялась) Результаты рентгеноструктурного анализа представлены на рисунке 3.

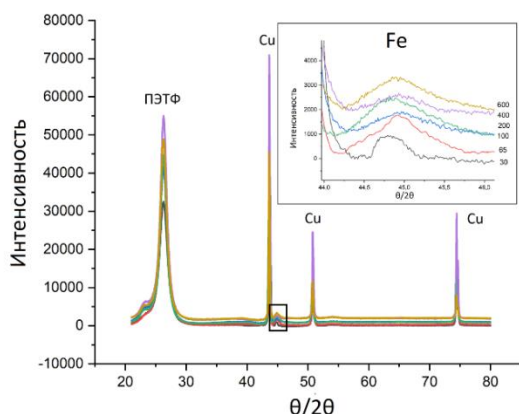


Рис. 3. Рентгеноструктурные данные для железных НП диаметрами 30, 65, 100, 200, 400 и 600 нм

Анализ полученных данных показал, что на всех образцах пики железа от НП имеют малую интенсивность, а наибольшей интенсивностью обладают пики меди от медной подложки.

Положение пиков железа от образца к образцу меняется, что говорит об изменении параметра решетки.

Для расчета межплоскостного расстояния используется закон Брэгга-Вульфа:

$$n\lambda = 2d\sin\theta, \quad (1)$$

где n – целое число, описывающее порядок дифракционного отражения; λ – длина волны; d – межплоскостное расстояние между отражающими плоскостями; θ – угол, который составляет падающий или дифрагированный луч с отражающей плоскостью. [3]

Расчет параметров решетки представлен в таблице 3.

Таблица 3. Расчет параметров решетки

ϕ , нм	d , Å	a , Å	2θ
30	2,026851	2,865	44,81
65	2,02129	2,858	44,94
100	2,020012	2,857	44,97
200	2,025994	2,865	44,83
400	2,021717	2,859	44,93
600	2,022571	2,860	44,91

Изменения носят систематичный характер за исключением образца с диаметром НП 200 нм. С повышением диаметра НП межплоскостное расстояние уменьшается. Наибольшее изменение наблюдается при переходе от образца с диаметрами 65 нм к образцу с диаметрами 30 нм.

Результаты показывают, что наблюдается уменьшение параметра решетки НП с увеличением их диаметра, однако величина этого изменения не значительна, что дает возможность предполагать, схожие структурные параметры НП разных диаметров.

У всех НП параметр решетки уменьшается по сравнению с объемным материалом ($a = 2,866$ Å). Это может быть связано с особенностями роста в условиях ограниченного объема, со «сдавливанием» растущего металла в порах. Слабая зависимость параметров решетки от диаметра НП может быть связана с изменениями других условий роста. Известно, что последние зависят от скорости диффузии и кислотности раствора. В свою очередь оба этих параметра зависят от диаметра порового канала.

Заключение

Были исследованы особенности заполнения пор матриц с разным диаметром. Проведен рентгеноструктурный анализ, который показал, что с увеличением диаметра НП межплоскостное расстояние уменьшается.

Благодарность

Измерения были выполнены на базе ФНИЦ Кристаллографии и Фотоники РАН.

Список литературы

1. Электроосаждение металлов группы железа в поры трековых матриц для получения нанопроволок/ Коротков В.В., Кудрявцев В.Н., Кругликов С.С., Загорский Д.Л., Сульянов С.Н., Бедин С.А.// Гальванотехника и обработка поверхности. 2015. Том 23, № 1. С.24-33.
2. John C. Hulteen and Charles R. Martin., A general template-based method for the preparation of nanomaterials, Department of Chemistry, Colorado State University, Fort Collins, CO 80523, USA
3. Цыбуля С.В. Лекции по рентгеноструктурному анализу: учеб.-методич. пособие / С.В. Цыбуля; Новосибирский государственный университет Институт катализа им. Г.К.Борескова СО РАН – Новосибирск, 2012. – 32 с.

НАНОКРИСТАЛЛИЧЕСКИЙ ОКСИД НИКЕЛЯ Р-ТИПА ДЛЯ ПЕРОВСКИТНОЙ ОПТОЭЛЕКТРОНИКИ С ПРИМЕНЕНИЕМ ЛУЧЕВЫХ ТЕХНОЛОГИЙ

Бронников О.А., Гостищев П.А.,
Лучников Л.О., Саранин Д.С.
«Национальный исследовательский
технологический университет "МИСИС"»,
Лаборатория перспективной
солнечной энергетики (L.A.S.E.)

Аннотация

Данное исследование посвящено изучению оптических, электрических и структурных свойств нанокристаллических слоёв окислов никеля, полученных ионно-лучевым методом. Также в работе рассмотрено применение NiO_x в качестве дырочно-транспортного слоя в перовскитных солнечных элементах. Было определено, что термический отжиг в атмосфере воздуха позволяет повысить прозрачность плёнок в видимом и ближнем инфракрасном диапазоне длин волн до 80% -85%, а также улучшить структурные и транспортные свойства плёнок.

Полученные перовскитные солнечные элементы достигли эффективности в 16,78 % при напряжении холостого хода, равном 1,028 В, и токе короткого замыкания 22,124 $\text{мА} \cdot \text{см}^{-2}$.

Введение

На сегодняшний день для разработки устройств тонкопленочной фотовольтаики и оптоэлектроники требуется удешевление стоимости технологических процессов. Перовскитные солнечные элементы — это р-і-п или п-і-р структуры, включающую в себя фотогенерирующий слой, электрон- и дырочно-транспортные слои [1]. Для их нанесения используются растворные и вакуумные методы (магнетронное напыление, электронно-лучевое напыление и т. д.).

В данной работе представлен новый подход получения нанокристаллических слоёв оксида никеля методом ионно-лучевого напыления, являющимся одной из перспективных альтернатив распространённому в полупроводниковой промышленности магнетронному распылению, благодаря отсутствию взаимодействия ионного луча с тонкопленочной структурой [2], формирующейся на подложке, а также избежанию перегрева будущих устройств [3].

Основными задачами работы является разработка оптимального технологического процесса получения тонких плёнок NiO_x с контролируемыми свойствами для применения в создании р-і-п солнечных элементов.

Создание плёнок NiO_x

Подложками для напыления оксида никеля являлись стёкла 25·25 мм с предварительно нанесённым прозрачным электродом ИТО. После ультразвуковой очистки подложки устанавливались в вакуумную камеру, где перед началом напыления поддерживается давление $4 \cdot 10^{-2}$ Па.

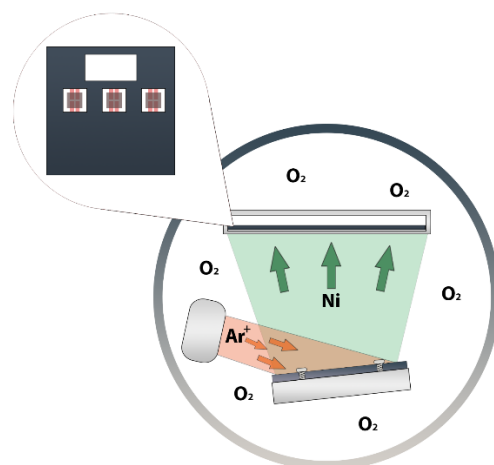


Рис. 1. Схематическое изображение процесса ионно-лучевого напыления NiO_x

Формирование плёнки происходит в процессе распыления атомов металлической мишени Ni при помощи ионов аргона, и осаждении атомов никеля на подложки, находящиеся в потоке кислорода, окисляющем плёнку. Во время проведения процесса напыления поддерживается напряжение на источнике ионов равное 1 кВ, а также величина ионного тока 150 мА. Данная конфигурация процесса позволяет получить скорость напыления NiO_x равной 1,2 нм/мин.

Далее плёнки были подвергнуты термическому отжигу в атмосфере воздуха. Так как слой ИТО при обработке температурой выше 350°C - 400°C начинает деградировать и его сопротивление вырастает в 2 и более раз [4], то отжиг структур Стекло/ИТО/ NiO_x проводился при 300°C в течение 1 часа.

Характеристики плёнок NiO_x

Оптоэлектрические свойства плёнок NiO_x в большей степени зависят от содержания в них кислорода и числа дефектов, таких как никелевые вакансии[5].

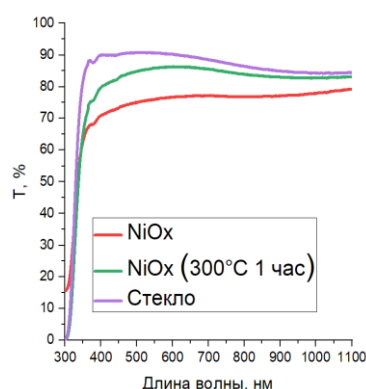


Рис. 2. Спектры пропускания плёнок оксида никеля напылённых на стекло

Таким образом, контролировать свойства плёнок можно с помощью термического отжига в атмосфере воздуха. Такая обработка улучшает прозрачность плёнки (рис. 2), которая составила не менее 80 % в интервале длин волн от 450 нм до 1100 нм и достигает 85 % при 550 нм - 600 нм. Оптическая ширина

запрещённой зоны E_g при отжиге выросла с 3,28 эВ до 3,77 эВ.

Результаты АСМ (атомно-силовой микроскопии) говорят об увеличении среднего размера зерна при отжиге 300°C до 14,2 нм, без отжига он составил 9,8 нм. Холловские измерения подвижности дырок показывают, что до и после отжига величина μ_p соответственно равна 4,05 $\text{см}^2\text{В}^{-1}\text{с}^{-1}$ и 52,56 $\text{см}^2\text{В}^{-1}\text{с}^{-1}$.

Так как за счёт отжига дефектов уменьшается число свободных дырок [6], наблюдается снижение удельной электропроводности σ с 434 $(\text{Ом}\cdot\text{см})^{-1}$ до 9 $(\text{Ом}\cdot\text{см})^{-1}$.

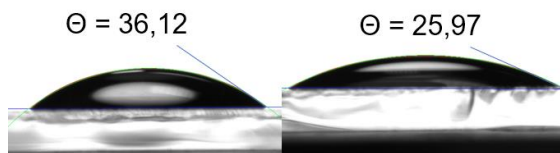


Рис. 3. Изменение краевого угла смачивания до (слева) и после (справа) термического отжига

Отожжённые при 300°C в течение 1 часа демонстрируют уменьшение краевого угла смачивания (рис. 3) и средней шероховатости R_a с 149,10 нм до 102,80 нм, что говорит об улучшении адгезии при нанесении перовскита на поверхность NiO_x .

Приборные характеристики

Данные плёнки были использованы при сборке перовскитных СЭ (солнечных элементов) p-i-n архитектуры, имеющих структуру ИТО/ NiO /CsFAPbI₃/C60/BCP/Cu (рис. 4) и активную площадь 0,14 см^2 . Толщина слоя NiO_x составила 7 нм.

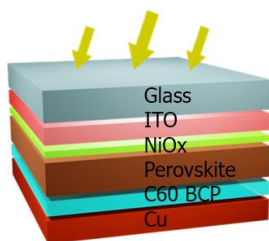


Рис. 4. Структура собранного СЭ

Лучшее устройство в данной сборке показало эффективность в 16,78%, напряжение холостого хода $U_{xx} = 1,028$ В и ток короткого замыкания $I_{кз} = 22,124 \text{ мА}\cdot\text{см}^{-2}$, что подтверждает успешность применения NiO_x в качестве селективно транспортного слоя.

На рисунке 5 изображена ВАХ (вольт-амперная характеристика) СЭ с NiO_x , на прямой ветви которой заметно значительное влияние последовательного сопротивления, что говорит об необходимости дальнейшей оптимизации процесса формирования слоя оксида никеля для СЭ.

Заключение

В данной работе был получен технологический маршрут создания нанокристаллического оксида никеля ионно-лучевым методом. Скорость напыления NiO_x составила 1,2 нм/мин.

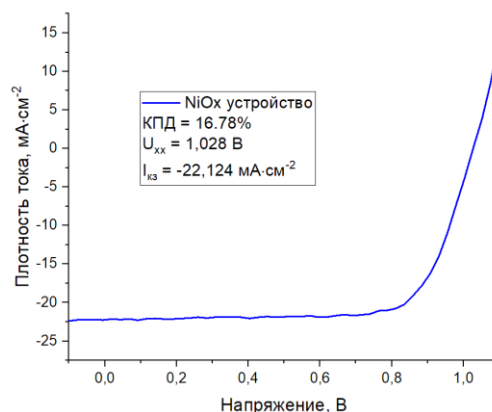


Рис. 5. Вольт-амперная характеристика лучшего устройства с NiO_x

Перовскитные солнечные элементы с р-транспортным слоем NiO_x , полученным с помощью ионно-лучевого напыления демонстрируют высокое напряжение холостого хода $U_{xx} = 1,028$ В, ток короткого замыкания $I_{кз} = 22,124 \text{ мА}\cdot\text{см}^{-2}$ и КПД = 16,78 %.

Также с целью оптимизации изготовления СЭ были исследованы изменения свойств плёнок NiO_x при отжиге на воздухе в течение 1 часа. Было определено, что термическая обработка положительно влияет на подвижность дырок (увеличение с 4,05 $\text{см}^2\text{В}^{-1}\text{с}^{-1}$ до 52,56 $\text{см}^2\text{В}^{-1}\text{с}^{-1}$) и размер зёрен (рост с 9,8 нм до 14,2 нм) тонкоплёночного оксида никеля, а также на его коэффициент пропускания, достигающий 85 % в интервале от 550 нм до 600 нм и 80 % в ближней инфракрасной области спектра (до 1100 нм).

Данная технология не требует использования дополнительного легирования, применения буферных слоёв, или других сложных модификаций, однако нуждается в дальнейшей оптимизации для получения более высоких приборных характеристик СЭ. Исследования в этом направлении могут значительно удешевить производство перовскитных СЭ.

Благодарность

Авторы благодарят финансовую поддержку гранта К2-2022-011 выполненного в рамках программы Приоритет 2030 в НИТУ МИСИС.

Список литературы

1. Yaghoobi Nia N. et al. Chapter Five - Perovskite solar cells // Solar Cells and Light Management / ed. Enrichi F., Righini G.C. Elsevier, 2020. P. 163–228.
2. Kuzmichev A. et al. Characteristics of flows of energetic atoms reflected from metal targets during ion bombardment // Journal of Physics: Conference Series. Institute of Physics Publishing, 2016. Vol. 729, № 1.
3. Ristau D., Gross T. Ion beam sputter coatings for laser technology // Advances in Optical Thin Films II. SPIE, 2005. Vol. 5963. P. 596313.
4. Zardetto V. et al. Substrates for flexible electronics: A practical investigation on the electrical, film flexibility, optical, temperature, and solvent resistance properties // J Polym Sci B Polym Phys. 2011. Vol. 49, № 9. P. 638–648.

5. Poulain R. et al. Electronic and Chemical Properties of Nickel Oxide Thin Films and the Intrinsic Defects Compensation Mechanism // ACS Appl Electron Mater. 2022. Vol. 4, № 6. P. 2718–2728.

6. Horak P. et al. Nickel oxide films by thermal annealing of ion-beam-sputtered Ni: Structure and electro-optical properties // Thin Solid Films. 2017. Vol. 640. P. 52–59.

МЕТОД ОПРЕДЕЛЕНИЯ ТЕМПЕРАТУРЫ ЭКСТРУЗИИ ПЛАСТИКА

*Бадалян Р. К., Кержеманкина К. В.,
Сафронов М. А., Тыщевский А. Ю., Эфендиев С. В.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент компьютерной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Ежегодно в России образуется около 3 млн т пластиковых отходов. На данный момент из них перерабатывается лишь 10–12%.

В данной работе описаны базовые принципы работы устройства по переработки пластиковых отходов, оставшихся после печати на 3D-принтере, сравнение его с существующими аналогами, моделирование тепловых процессов печатной платы и метод выбора температурного алгоритма.

В качестве результата проделанной работы были получены сведения о температуре, необходимой для плавки пластика и сохранения всех его физико-химических свойств, была выбрана элементная база печатной платы и был подобран наилучший алгоритм для преобразования сопротивления в значение температуры.

Введение

Целью работы является выбор оптимального метода определения температуры расплавленного материала. Для достижения этой цели необходимо выполнить поставленные задачи:

1. Провести моделирование тепловых процессов
2. Выполнить тепловое моделирование печатной платы и определить элементную базу печатной платы
3. Провести исследование выбранного датчика температуры
4. На основании полученных данных подобрать алгоритм для вывода температуры

Проблема

Ежегодно в России образуется около 3 млн т пластиковых отходов. Специалисты НИУ ВШЭ провели анализ данных в 2017 году и пришли к выводу, что на переработку от этого количества отправляется всего 10–12%. Система раздельного сбора отходов в России пока не позволяет добиться полной загрузки профильных предприятий по переработке пластика. Из-за этого огромное количество пластиковых изделий оказываются на полигонах и свалках. Синтети-

ческий материал начинает разлагаться, в процессе чего выделяются вредные вещества: углекислый и угарный газы, фосген и бромистый водород [1].

В ходе работы 3D-принтера также образуются пластиковые отходы, которые представляют из себя поддержки печати (вспомогательный материал, используемый в 3D-печати для построения сложных объектов и увеличения качества и стабильности построения) или бракованные изделия, поэтому пользователю приходится выкидывать около 20% пластика, не получая с этого выгоду. ReCoil System позволяет перерабатывать пластиковые отходы, в следствии чего затраты на материал будут снижаться, а в условиях работы с большим количеством 3D-принтеров интеграция данного устройства приведет к минимизации отходов при производстве.

Принцип работы ReCoil System

В камеру устройства засыпаются пластиковые гранулы, которые после плавления кольцевым нагревателем при помощи шнека выдавливаются через сопло заданного диаметра в нить для 3D-печати. Полученная нить равномерно наматывается на вращающуюся катушку, так как в таком виде хранится и эксплуатируется пластик для 3D-принтеров. Этот процесс возможен благодаря синхронизации шаговых двигателей, отвечающих за одновременную экструзию и намотку пластика. Все электронные и управляющие модули будут располагаться в корпусе печатной платы.

Обзор существующих аналогов

Одним из аналогов разрабатываемой системы является экструдер компании Wellzoom. Данный экструдер способен перерабатывать ABS, PLA и PVA пластики и производить нить диаметром 1,75 и 3 мм с производительностью от 300 до 650 мм/мин благодаря возможности выбора одного из трех режимов работы. Максимальная рабочая температура экструдера составляет 300 °С. Существенный недостаток такой системы - пользователь не может задать скорость экструзии, из-за чего отсутствует возможность переработки альтернативных видов пластика.

В качестве еще одного аналога можно рассмотреть экструдер Filabot Original американской компании Filabot. Устройство предназначено для работы с ABS, PLA и HIPS пластиками. В качестве расходного материала используется гранулированный пластик, но возможно и использование пластиковых отходов при том условии, что размер частиц не превышает 5мм [2]. Данный экструдер также производит нить диаметром 1,75 и 3 мм, при этом есть функция контроля температуры от 40 до 400 °С, что позволяет системе производить качественную нить из широкого спектра материалов, включая композиты с древесным наполнителем. Но и у этой разработки есть несколько недостатков: во-первых, скорость вращения шнека составляет 35 оборотов в минуту. Такое решение позволяет перерабатывать материалы с различной пластичностью, что снижает риск обрыва нити при наматывании на катушку. Однако низкая скорость экструзии сказывается на производительно-

сти – данное устройство способно производить нить со скоростью 250 мм/мин. Во-вторых, рассматриваемый экструдер не оборудован катушкой для наматывания нити.

В ReCoil System предусмотрена катушка для автоматического накручивания нити, при этом шаговый двигатель, отвечающий за этот процесс синхронизирован с двигателем, отвечающим за вращение шнека, так как они оба управляются одним контроллером. Также при эксплуатации устройства пользователь может самостоятельно выбирать необходимую скорость экструзии в диапазоне от 20 до 60 об/мин и устанавливать необходимую температуру для плавления пластика (таблица 1). Благодаря такой функции можно перерабатывать альтернативные виды пластика, изначально непредназначенные для 3D-печати.

Таблица 1. Сравнение экструдеров

	ReCoil System	Wellzoom	Filabot
Размеры	-	502x138x252 мм	431x178x203 мм
Диаметр прутка	1,75 и 3 мм	1,75 и 3 мм	1,75 и 3 мм
Скорость вращения шнека	20-60 об/мин	-	35 об/мин
Контроль температуры	✓	✗	✓
Контроль скорости	✓	✓	✗

Моделирование тепловых процессов

Самыми важными этапами разработки устройства являются моделирование тепловых процессов и моделирование печатной платы. От данных, полученных при моделировании, зависит множество факторов:

- тип датчика, применяемый для определения температуры;
- размеры шнека и экструдера;
- вся элементная база печатной платы;
- система охлаждения;
- размеры будущей конструкции.

После создания прототипа устройства и тестирования процесса экструзии необходимо разработать 3D-модель будущей конструкции, затем провести моделирование тепловых процессов в среде SolidWorks. Это необходимо для определения температуры нагрева камеры и шнека. В зависимости от полученных данных, в программу управления экструдером будут внесены соответствующие значения температуры для каждого типа пластика (при моделировании также необходимо добавить слой термоизоляции и установить значение температуры окружающей среды). На этапе теплового моделирования также решается вопрос о необходимости внедрения системы охлаждения для пластика и корпуса блока управления. Данный шаг также необходим, ведь при слишком высокой температуре пластиковая нить будет неравномерно накручиваться на катушку, из-за чего может пострадать 3D-печать из этой нити.

При разработке устройства было принято решение интегрировать кулеры для охлаждения катушки и корпуса печатной платы, так как она расположена под экструдером, где выделяется большое количество тепла. Параллельно с моделированием тепловых процессов отдельно проводится тепловое моделирование печатной платы. От температуры компонентов

зависит надежность, долговечность и работоспособность всей печатной платы [3]. Так при разработке принципиальной схемы необходимо рассчитать выделяемую на каждом компоненте мощность, внести эти значения в SolidWorks, задать материал компонента и выбрать способ передачи тепла. Таким образом при разработке печатной платы были получены следующие результаты (см. рис. 1): драйверы шагового двигателя a4988 нагрелись до 63°C. Поэтому было принято решение размесить радиатор и кулер для охлаждения драйвера.

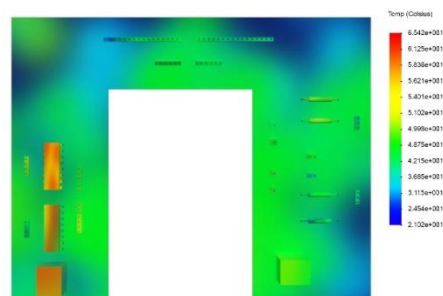


Рис. 1. Тепловое моделирование печатной платы

Анализ температурных алгоритмов

После проведения теплового моделирования и выбора элементной базы необходимо выбрать способ определения и вывода температуры на дисплей. Так как в разрабатываемом устройстве в качестве датчика температуры применяется NTC-термистор [4], сопротивление которого уменьшается с увеличением температуры, было принято решение провести исследование термистора. Для этого термистор подключался к плате Arduino через резистор 100 кОм, затем на монитор выводилось значение сопротивления термистора в зависимости от температуры окружающей среды (таблица 2):

Таблица 2. Экспериментальные данные расчета зависимости сопротивления термистора от его температуры

Температура (°C)	Сопротивление (кОм)
24	2,460
25	2,451
26	2,449
27	2,447
28	2,445
29	2,442
30	2,439
31	2,436
32	2,434
34	2,432
36	2,429
38	2,426
41	2,424
44	2,422
47	2,410
50	2,408
53	2,400
60	2,397
100	2,392

Затем на основании полученных данных был построен график (см. рис. 2). График экспоненциально убывает – значит термистор работает корректно, однако использовать выведенную формулу, полученную на основе данных из таблицы 2 нельзя, так как при подключении термистора в печатную плату его сопротивление и посчитанные значения могут измениться, и, как следствие, применится неверная формула для конкретного термистора. Более того при каждой замене термистора или платы появляется необходимость заново вывести формулу и обновлять скетч управления [5].

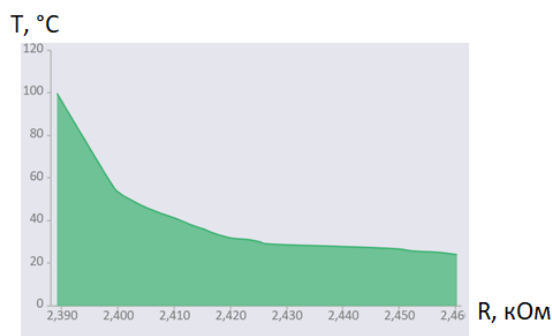


Рис. 2. График зависимости сопротивления термистора от его температуры

Вторым способом подсчета и вывода температуры является использование уже готового уравнения для описания NTC-термистора. Самым подходящим в данной ситуации является алгоритм Стейнхарта-Харта (SHE) (1), так как он имеет такую же экспоненциальную зависимость [6, 7].

$$\frac{1}{T} = A + B \ln(R_{th}) + C (\ln(R_{th}))^3, \quad (1)$$

где T – температура, A , B , C – константы, R_{th} – сопротивление терморезистора при температуре T

Для использования этого алгоритма необходимо рассчитать константы A , B и C . То есть необходимо выполнить три измерения, что также может привести к образованию ошибок при расчетах. Более того, эти измерения необходимо выполнять повторно каждый раз при замене термистора или печатной платы. Вместо этого, есть более простое уравнение, которое менее точно, но содержит только одну константу. Эта константа обозначена как β , и заявляется производителем в технической документации термистора (2):

$$R_{th} = R_0 e^{\beta \left(\frac{1}{T} + \frac{1}{T_0} \right)}, \quad (2)$$

где β – константа, R_0 – сопротивление при контрольной температуре T_0

Подставив формулу в программное обеспечение ReCoil System, был проведен тест экструзии PLA и ABS пластика. Оптимальная температура экструзии PLA пластика, при которой у готовой нити сохраняется изначальная прочность, гибкость и не разрушается цветовой пигмент, составляет 150 °C (см. рис. 3), а температура экструзии ABS пластика – 190 °C (см. рис. 4).



Рис. 3. Расчетная температура экструзии PLA пластика

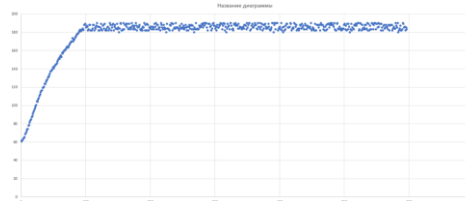


Рис. 4. Расчетная температура экструзии PLA пластика

Заключение

В результате моделирования тепловых процессов экструдера были получены сведения о температуре, необходимой для плавки определенного вида пластика и сохранения его прочности, гибкости и цветового пигмента. Также на основании полученных данных была выбрана элементная база печатной платы. На этапе исследования термистора был подобран наилучший алгоритм для преобразования сопротивления в значение температуры. В дальнейшем полученные данные понадобятся для определения корректной скорости экструзии определенного типа пластика.

Список литературы

1. Как устроена сфера переработки пластика в России. – URL: <https://news.solidwaste.ru/2022/07/kak-ustroena-sfera-pererabotki-plastika-v-rossii/>
2. Портативные экструдеры прутка. Изготовление филамента. Как самому сделать расходные материалы для 3D-печати. – URL: https://3dtoday.ru/wiki/made_filament
3. Расчет ширины дорожки печатной платы в зависимости от силы тока. – URL: <https://radioprogram.ru/calculator/1>
4. Шашков, А.Г. Терморезисторы и их применение. – URL: <https://specknigi.ru/termistory-cto-eto-takoe/>
5. Бадалян Р. К., Сафронов М. А., Эфендиев С. В. Разработка устройства для переработки отходов пластика после 3D-печати: Выпускная квалификационная работа по направлению подготовки 09.03.01 образовательная программа «Информатика и вычислительная техника» - 2022. - С. 31–33.
6. Rana, K., Kumar, V., Kumar, A., Chandel, A., Kataria, A. FPGA Implementation of Steinhart–Hart Equation for Accurate Thermistor Linearization // “IEEE SENSORS” journal - 2018. - Vol. 18, No. 6, P. 1–2.
7. Mahani, R., Abdel, D. A Thermistor with Variable Rate of Negative Temperature Coefficient of Resistance Made from Egyptian Raw Materials // Journal of Electronic Materials – 2020. - Vol. 49, No. 12 – P. 3.

Секция «Информационная безопасность»

ПРИМЕНЕНИЕ ПЛАТФОРМЫ ТХМ ДЛЯ АНАЛИЗА ТЕКСТОВ РАЗЛИЧНОГО ТИПА

Бурба А.В.¹, Фокина А.И.²

*¹Российский университет дружбы народов,
факультете физико-математических
и естественных наук,
математический институт имени
С.М. Никольского,*

*²Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ*

Аннотация

В работе рассматриваются результаты исследования текстов на основе метода анализа соответствий платформы корпусного анализа ТХМ. Исследованы не связанные между собой корпуса противоправных и поэтических текстов для проверки применимости и эффективности методики на несходных наборах текстов. В результате выявлена результативность применения анализа соответствий для корпусов различного типа. Сделан вывод о возможности применения данного инструмента платформы ТХМ для оценки качества составления какого-либо корпуса и его дальнейшего применения в качестве обучающей выборки для решения задач машинного обучения.

Введение

Как в теоретических исследованиях, так и практических задачах сегодня активно применяется автоматизированный анализ текстов на естественных языках. В числе ярких примеров областей его применения: поиск информации, электронный документооборот, взаимодействие с помощью электронных онлайн средств посредством текстовых сообщений, анализ данных, классификация текстов, построение языковых моделей и другие.

В то же время развитие средств онлайн общения и взаимодействия делает актуальной проблему распространения экстремизма в сети. Из этого следует востребованность разработки методик и программных инструментов интеллектуального анализа текстов на естественных языках для определения среди них экстремистского и террористического смысла.

С помощью анализа поэтических словарей можно определить влияние языковых особенностей и исторических фактов на тексты. Политические, экономические и социальные события оказывают весомый эффект на становление литературного языка и популярность определенных тем в произведениях авторов.

Платформа ТХМ и дополнительная функциональность

В данной работе анализ основан на методике анализа соответствий программного комплекса – плат

формы ТХМ [1], которая обладает различными средствами для исследования корпусов текстов.

Методика выделения псевдооснов словоупотреблений [2-4], которая в данном исследовании применяется во время предобработки в момент загрузки в платформу ТХМ, расширяет потенциал корпусного анализа.

Зачастую комментарии и сообщения в сети, а иногда и поэтические тексты, обладают большим количеством неологизмов и жаргонизмов, отсутствующих в стандартных словарях. Аналитический метод выделения псевдооснов дает возможность обрабатывать подобные тексты и определять их дифференцирующие признаки.

Метод структурных схем [5-6] является основой инструмента выделения псевдооснов. Методика состоит в их получении путем анализа словоизменительных аффиксов словоформ.

Анализ соответствий

Настоящая работа опирается на анализ соответствий – метод анализа корпуса текстов, который разделен на части, называемые подкорпусами, исходя из определенного признака. В основе инструмента лежит механизм определения независимости категориальных переменных. В процессе решения данной задачи методом анализа соответствий исследуется частота совместного появления значений данных переменных. Результат достигается с помощью оценки расположения набора точек, относящихся к одной из групп, на плоскости. Исследование графического представления анализа соответствий дает возможность оценки взаимного расположения и упорядоченности подкорпусов, а также отдельных словоформ, входящих в их состав.

Ж.-П. Бензекри [7] является автором методики анализа соответствий. Она является одним из инструментов платформы ТХМ, имплементирована в пакете FactoMineR для платформы R [8].

Исследуемые корпуса текстов

В состав корпуса противоправных текстов, который анализировался в ходе исследования, вошли тексты экстремистской направленности, а также группа сходных по тематике, но в то же время нейтральных по стилю текстов [9-10]. В последнюю включены тексты с оппозиционных и проправительственных политических блогов, разрешенные тексты религиозного содержания, новостные статьи. Корпус противоправных текстов был собран вручную экспертами и содержит примерно 3,3 миллиона словоупотреблений.

Корпус включает в себя экстремистские тексты семи категорий: терроризм, идеологические тексты, религиозная ненависть, сепаратизм, национализм, агрессия и призывы к беспорядкам, фашизм, а также нейтральные тексты.

Подкорпус «Терроризм» включает в себя тексты запрещенных организаций в РФ с обращениями участников данных группировок, а также идеологическую пропаганду. «Идеология» состоит из текстов, утверждающих превосходство определенной религии, распространяющих ложные трактовки священных писаний и призывы к обращению в иную веру. К данному подкорпусу близок подкорпус «Религия», однако, последний включает в себя более активные призывы к действиям против других религий, распространяются недостоверные сведения, которые формируют отрицательный образ определенных конфессий. «Сепаратизм» содержит оскорбления и угрозы в адрес определенных этнических групп, концепцию отделения субъектов. Тексты о враждебной настроенности к этническим группам, их уничтожении и ограничении вошли в состав подкорпуса «Национализм». «Агрессия» включает негативные тексты по отношению к власти, призывы к митингам. Подкорпус «Фашизм» содержит идеи данной идеологии, а также неонацизма и геноцида.

Также в ходе исследования был проведен анализ корпуса текстов поэзии серебряного века. Рассмотрение результатов анализа для независимых корпусов, как в данном случае корпуса противоправных текстов и корпуса поэзии, позволяет осуществить сопоставление эффективности применения методики для корпусов различных типов.

В качестве поэтического корпуса было рассмотрено четыре литературных направления конца XIX – начала XX века. Серебряный век русской поэзии включает в себя несколько самобытных и ярких литературных течений, основные из них были рассмотрены в качестве подкорпусов.

В корпус поэзии входит сборник поэзии из около 101 тыс. словоупотреблений. Отбор текстов производился вручную: было выбрано четыре литературных течения (акмеизм, символизм, футуризм, имажинизм), представленные в корпусе в примерно равном объеме. В каждом из течений рассматривалось по четыре автора, а внутри подкорпусов стихотворения были разделены по годам творчества создателей в период развития того или иного литературного течения.

Корпусной анализ

Анализ соответствий отражает пространственное расположение подкорпусов на основе анализа частот совместного появления значений переменных.

На рис. 1, где представлены результаты анализа соответствий по псевдоосновам для отдельных слов корпуса противоправных текстов, «Нейтральный» подкорпус находится в отдалении от каждого из экстремистских на первой (горизонтальной) оси. «Идеология», «Религия», «Терроризм», «Национализм», «Агрессия» и «Сепаратизм» близки по своим показателям и образуют «облако» в первой четверти координатной плоскости.

Подкорпус «Фашизм», находящийся в четвертой четверти, близок ко всем экстремистским по 1-й оси (46% вариации) и противостоит им по 2-й оси (13% вариации).

Стоит отметить близость подкорпуса «Нейтральный» к точке пересечения координатных осей с уклоном в отрицательные значения по оси абсцисс. В то же время все противоправные подкорпуса обладают положительными значениями по первой оси.

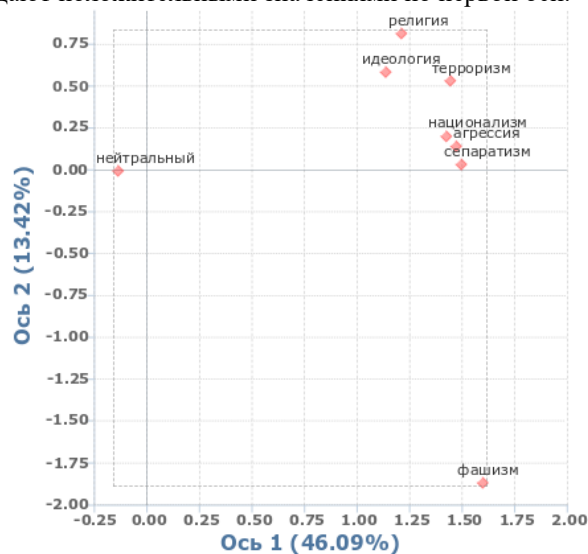


Рис. 1. Анализ соответствий для противоправного корпуса

На рисунке 2 представлен график анализа соответствий для начальных форм слов направлений поэзии серебряного века в осях 1 и 2. Ось 1 характеризует хронологию возникновения течений. Так, первым из рассматриваемых направлений был символизм, за которым в 1910 году последовали акмеизм и футуризм, далее сформировался имажинизм. При этом футуризм охватил период и акмеизма, и имажинизма, что на графике отображено его пограничным расположением по оси 1 между соответствующими точками.

По оси 2 практически совпадают подкорпусы символизма и акмеизма с координатами 0,07 и 0,08 соответственно, что объясняется возникновением акмеизма из символизма. При этом были внедрены новые идеи ясности образов и отрицания мистического, что на графике характеризуется расположением точек подкорпусов по разные стороны от оси 2.

Стоит отметить, что подкорпусы «Футуризм» и «Имажинизм» по оси 1 находятся недалеко друг от друга, эта особенность вызвана тем, что исторически имажинизм можно назвать литературным направлением, которое берёт своё начало из футуризма, следовательно, подкорпусы имеют схожие особенности, однако у «Футуризма» они выделены более явно. Интересно то, что несмотря на наибольшую близость этих подкорпусов по оси 1, они являются абсолютно противоположными по отношению друг к другу по оси 2, что говорит нам о самобытности и независимости этих течений.

Рассматривая «Акмеизм», можно говорить о том, что сравнительно недалёкое расположение по оси 1 от футуризма и имажинизма связано с совпадающими годами развития этих течений.



Рис. 2. Анализ соответствий для течений поэтического корпуса

Заключение

Применение анализа соответствий платформы TXM является результативным при исследовании и разделении текстов различных направленностей. Как для противоправных, так и для поэтических текстов, возможно применение методов корпусной лингвистики и выделение дифференцирующих признаков для задач различных типов: определение как тематики противоправных текстов, так и поэтических стилей. Таким образом, с помощью анализа соответствий платформы TXM возможно исследование текстов внутри заданной тематики корпуса, а также оценить качество классификации текстов в корпусе и возможность его дальнейшего применения в качестве обучающей выборки для задач машинного обучения.

Список литературы

1. Heiden S. The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme // 24th Pacific Asia Conference on Language, Information and Computation - PACLIC24 / Ed. R. Otoguro, K. Ishikawa, H. Umemoto, K. Yoshimoto and Y. Harada. Institute for Digital Enhancement of Cognitive Development, Waseda University, Sendai, Japan. 2010. P. 389—398. URL: <http://halshs.archives-ouvertes.fr/halshs-00549764>
2. Лаврентьев А. М., Соловьев Ф. Н., Суворова М. И., Фокина А. И., Чеповский А. М. Новый комплекс инструментов автоматической обработки текста для платформы TXM и его апробация на корпусе для анализа экстремистских текстов // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2018. Т. 16, № 3. С. 19–31.
3. Chepovskiy A., Devyatkin D., Smirnov I., Ananyeva M., Kobozeva M., Solovyev F. Exploring linguistic features for extremist texts detection (on the material of Russian-speaking illegal texts). In: 2017 IEEE International Conference on Intelligence and Security Informatics: Security and Big Data, ISI 2017. Institute of Electrical and Electronics Engineers Inc., 2017, p. 188–190.

4. Соловьев Ф. Н. Автоматическая обработка текстов на основе платформы TXM с учетом анализа структурных единиц текста // Вестник НГУ. Серия: Информационные технологии, 2020. Т. 18, №1. С. 74–82.

5. Egorova E., Chepovskiy A., Lavrentiev A. A structural pattern based method for automated morphological analysis of word forms in a natural language. Journal of Mathematical Sciences, 2016, vol. 214, no. 6, p. 802–813.

6. Болховитянов А. В., Чеповский А. М. Методы автоматического анализа словоформ // Информационные технологии. 2011. № 4 (176). С. 24–29.

7. Benzécri J.-P. L'analyse des données: l'analyse des correspondances. 2nd ed. Paris: Dunod, 1979. Vol. 2.

8. Lê S., Josse J., & Husson F. FactoMineR: an R package for multivariate analysis // Journal of Statistical Software. 2008. № 25 (1) P. 1–18.

9. Ананьева М. И., Кобозева М. В., Соловьев Ф. Н., Поляков И. В., Чеповский А. М. О проблеме выявления экстремистской направленности в текстах // Вестник Новосибирского государственного университета. Серия: Информационные технологии. Новосибирск. Федеральное государственное автономное образовательное учреждение высшего образования "Новосибирский национальный исследовательский государственный университет". 2016. Т. 14. № 4. С. 5–13.

10. Ананьева М. И., Девяткин Д. А., Кобозева М. В., Смирнов И. В., Соловьев Ф. Н., Чеповский А. М. Исследование характеристик текстов противоправного содержания // Труды Института системного анализа Российской академии наук. Москва. Федеральное государственное учреждение "Федеральный исследовательский центр "Информатика и управление" Российской академии наук. 2017. Т. 67. № 3. С. 86–97.

ВЫДЕЛЕНИЕ НЕЯВНЫХ ПЕРЕСЕКАЮЩИХСЯ СООБЩЕСТВ НА ГРАФЕ ВЗАИМОДЕЙСТВИЯ TELEGRAM-КАНАЛОВ

Аванесян Н.Л.¹, Зенькова В.В.², Попов В.А.¹

¹Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

²Московский физико-технический институт
(национальный исследовательский университет),
кафедра системных исследований

Аннотация

В работе рассматривается выделение неявных сообществ на графе взаимодействующих объектов, полученном при импорте сети каналов из мессенджера Telegram. Для оценки качества разбиения проведен психолингвистический анализ текстов полученных сообществ, а также выделены закономерности в зависимости от тематической направленности каналов. Представленное сочетание алгоритмическо-

го подхода для выделения сообществ и психолингвистического анализа соответствующих текстов имеет практическое применение для задач анализа информационного воздействия. К примеру, выявление текстов противоправной направленности.

Введение

В последнее время наблюдается стремительное распространение противоправной информации в сети Интернет, в частности, в социальных сетях. Экстремисты активно пропагандируют свои радикальные идеи и стараются повлиять на разум как можно большего количества людей (в основном, молодежи). Поэтому следует своевременно выявлять и блокировать такого рода информацию. А для этого необходимо разрабатывать методики анализа социальных сетей с целью выявления сообществ, анализа профилей пользователей и текстовой информацией, которой они обмениваются [1,2].

В том числе для решения этих задач в настоящее время активно исследуется тема построения и анализа графов взаимодействующих объектов, в частности сформированных на данных из социальных сетей, где вершинами являются пользователи или иные объекты, а ребрами – взаимодействия между ними [3-7]. Но помимо классических социальных сетей в современном мире активно набирают популярность кроссплатформенные системы обмена сообщениями, например, мессенджер Telegram. Помимо обмена личных сообщений между пользователями, многие частные лица, организации и корпорации широко используют Telegram как площадку для ведения своих публичных каналов. При этом Telegram предоставляет широкий спектр инструментов для авторов. В частности, автор в своем канале помимо уникального контента может сделать репост с другого канала, процитировать или упомянуть другие каналы, сделать ссылку на внешний URL. Все это возможности позволяют рассматривать множество Telegram-каналов и связи между ними как взвешенный граф взаимодействующих объектов.

Одними из методов анализа таких сетей является подход по выявлению активно взаимодействующих между собой объектов путем выделения неявных сообществ на графах [8-11] и подход подсчета центральности узлов графа для выявления вершин, которые играют решающую роль в сложных сетях [12, 13]. Классические меры центральности рассчитываются на основе свойств отдельно выбранной вершины, не учитывая структуру графа в целом, в том числе наличие неявных сообществ, что может быть не применимо для реальных сетей. Поэтому в [12] предлагается новая мера центральности, учитывающая разбиение графов на сообщества, а в [13] сравнивается 7 мер центральности на 15 реальных графах (часть из них получена из социальных сетей) для определения метрик, которые можно использовать на практике.

В данной работе рассмотрены графы взаимодействующих объектов, полученные из Telegram, описан и применен метод выделения неявных сообществ и проведен анализ текстов полученных сообществ.

Стоит отметить, что Telegram-каналы могут относиться к нескольким неявным сообществам сети, так как многие из них рассчитаны на широкую аудиторию, могут не иметь единую тематику постов и могут отражать различные взгляды и интересы. Это следует учитывать при разбиении графов на неявные сообщества и при их дальнейшем анализе.

Импорт данных из мессенджера Telegram

Для таких графов необходимо определиться с методикой построения ребер между вершинами, роль которых играют сами Telegram-каналы. В мессенджере Telegram были выделены три вида взаимоотношений между каналами: наличие общих внешних URL, упоминания и репосты. Учитывая эти особенности, ранее была разработана (U, M, R)-модель информационного взаимодействия [14].

На основе этой модели после импорта при помощи API (Application Programming Interface) реальных данных из Telegram строится взвешенный граф, в котором вершинами являются Telegram-каналы, а наличие ребер между ними и веса определяются имевшими место взаимодействиями.

Пусть $G(V, E)$ – граф, у которого V – множество Telegram-каналов, а E – множество всех возможных ребер – взаимодействий между парами каналов. На данном множестве ребер E определим весовую функцию $w(e_{AB})$ ($e_{AB} \in E; A, B \in V$) следующим образом:

$$w(e_{AB}) = U \times \delta_{e_{AB}}^U + M \times \delta_{e_{AB}}^M + R \times \delta_{e_{AB}}^R,$$

где $\delta_{e_{AB}}^U$ – количество общих уникальных внешних ссылок (URL) в постах у каналов A и B ; $\delta_{e_{AB}}^M$ – количество постов, где канал A упоминает канал B и наоборот B упоминает A ; $\delta_{e_{AB}}^R$ – количество репостов канала A у канала B , плюс наоборот B у A . Все указанные взаимоотношения рассматриваются за заранее выбранный период времени.

Для каждой пары Telegram-каналов из числа вершин графа V вычисляется указанная весовая функция w . Если она больше нуля, то между вершинами формируется ребро с посчитанным весом. Таким образом, на основе множества Telegram-каналов V и значений весовой функции w мы формируем итоговый взвешенный граф взаимодействующих объектов G . Данный граф сохраняется в специальный формат AVS, в котором хранятся атрибуты каждой вершины и ребра.

Также на данном этапе для каждого Telegram-канала скачиваются тексты его постов за выбранный период времени. Впоследствии это позволяет изучать не только структуру самих графов, но и содержание его объектов.

Алгоритм выделения неявных сообществ

Для выделения неявных сообществ на взвешенных графах взаимодействующих объектов сети Telegram был применен «метод Галактик» [15].

Данный метод основан на последовательном выделении пересекающихся сообществ на исходном

взвешенном графе, дальнейшем построении нового графа, в котором вершинами являются выделенные на первом шаге сообщества, называемые авторами «метавершинами». Взвешенные ребра нового графа между «метавершинами» строятся исходя из весов между каждой парой вершин исходного графа. Далее на новом графе выделяются непересекающиеся сообщества. В итоге получается разбиение исходного графа на пересекающиеся сообщества.

Выделение неявных сообществ на реальных данных мессенджера Telegram

Из мессенджера Telegram были импортированы данные о взаимодействии каналов и построены три графа взаимодействующих объектов.

Для всех графов применялись единые коэффициенты (1, 2, 3) в (U, M, R)-модели и рассматривался один и тот же промежуток времени с 8.12.2021 по 22.12.2021, за который были скачаны взаимодействия между каналами и тексты постов.

Таким образом было получено три графа:

- *G1* с начальной вершиной *@infernai_money* и глубиной скачивания, равной 3;
- *G2* с начальной вершиной *@sportsru* и глубиной скачивания, равной 5;
- *G3* с начальной вершиной *@stranavozmojnostey* и глубиной скачивания, равной 3.

Далее по методу Галактик на всех графах были выделены неявные сообщества. Для графов *G1*, *G2*, *G3* получено 17, 20 и 21 пересекающихся сообществ соответственно. На основе экспертных оценок были определены основные тематики и направленность выделенных сообществ. После этого для дальнейшего анализа для каждого выделенного сообщества были объединены тексты постов входящих в них Telegram-каналов за выбранный промежуток времени (8.12.21-22.12.21).

Психолингвистический анализ текстов выделенных сообществ

Для проверки эффективности выделения неявных сообществ на графе методом Галактик для каждого из рассматриваемых графов были посчитаны психолингвистические характеристики текстов постов этих сообществ и графа в целом. Данные психолингвистические характеристики определялись с помощью процедур автоматизированной обработки текстов на естественных языках, которые были описаны в [16-21].

В текстах для выделяемых словоупотреблений проводился автоматический морфологический анализ. Одной из лингвистических характеристик текста стала псевдооснова слова [22]. Именные и глагольные группы выделялись при помощи процедур синтаксического анализа [17, 18].

Для каждого текста методами компьютерной лингвистики [16, 17] были вычислены статистические характеристики текстов как возможные психолингвистические показатели. По результатам исследования было рассмотрено 23 показателя, разбитых на три различных группы: определяющие общие характеристики текстов (средние количества лексиче-

ских единиц); показывающие лексические характеристики текста (лексическое разнообразие); указывающие на использование синтаксических связей в словосочетаниях (относительные длины и составы именных и глагольных групп).

Объемы исследуемых текстов различных выделенных сообществ для графа *G1* варьировались от 0.5 Мб до 7 Мб, для графа *G2* – от 0.2 Мб до 9 Мб, а для графа *G3* – от 0.2 Мб до 3.5 Мб.

Результаты анализа сообществ

По результатам экспертных оценок на основе тематического содержания каналов полученные неявные сообщества объединялись по группам. К примеру, для графа *G1* были выделены следующие тематики: финансы/недвижимость, политика/новости, мода/искусство и другие. Для графа *G2*: спорт, политика/экономика, музыка и другие. В некоторых сообществах сложно выделить единую тематику, и они определяются как сообщества со смешанной тематикой. Важно отметить, что выделенные для исследуемых графов сообщества отличаются в большинстве случаев по психолингвистическим характеристикам своих текстов. Причем эти отличия коррелируют с экспертными оценками тематик выделенных сообществ.

С учетом вычисленных факторов для таких объединенных текстов можно выделить несколько (пять из 23) психолингвистических характеристик, которые наиболее явно изменяются в зависимости от рассматриваемого сообщества:

- Коэффициент лексического разнообразия 2 (ЛР2) – коэффициент разнообразия по псевдоосновам – отношение числа уникальных псевдооснов к числу словоупотреблений.
- Коэффициент действия 2 (КД2) – отношение количества глаголов и глагольных форм (деепричастий и причастий) к количеству прилагательных.
- Коэффициент опредмеченности действия (КОД) – соотношение количества глаголов (деепричастия и причастия исключаются) к количеству существительных.
- Коэффициент логической связности 1 (ЛС1) – отношение общего количества служебных слов (союзов и предлогов) к общему количеству предложений.
- Коэффициент связности лексики (СЛ) – отношение числа существительных и глаголов (деепричастия и причастия исключаются) к количеству прилагательных и наречий.

Значения указанных психолингвистических коэффициентов ЛР2, КД2, КОД, ЛС1 и СЛ представлены в таблицах результатов для выделенных сообществ №№ 2, 3, 5, 13, 14, 15 графа *G1*, которые по экспертной оценке являются сообществами по тематике «политика/новости» (Таблица 1) и для выделенных сообществ №№ 4, 9, 13, 16, 18 графа *G2*, которые по экспертной оценке являются сообществами по тематике «политика/экономика» (Таблица 2).

Из приведенных в таблицах (Таблица 1 и Таблица 2) результатов расчета видно, что для разных сооб-

ществ даже близких тематик обсуждения мы наблюдаем различия в значениях указанных психолингвистических характеристик. А это уже указывает на различия в лексике и семантике текстов сообществ (ЛР2, ЛС1 и СЛ) и на различия в степени побуждения к активности и действиям со стороны каналов из различных сообществ (КД2 и КОД).

Отметим, что указанные отличия по психолингвистическим характеристикам текстов наблюдаются для сообществ, которые относятся к одной тематической группе каналов. Что, в свою очередь, доказывает справедливость разделения этих каналов на разные сообщества. При том, что само выделение неявных сообществ осуществлялось исходя из структуры графа без учета специфики текстов. Как следствие, получено дополнительное подтверждение эффективности метода Галактик.

Таблица 1

Коэффициент	Номер сообщества					
	2	3	5	13	14	15
ЛР2	0,09	0,16	0,25	0,14	0,16	0,14
КД2	1,46	1,39	1,88	1,45	1,59	1,36
КОД	0,37	0,36	0,47	0,35	0,42	0,36
ЛС1	2,44	2,18	2,04	2,39	2,39	2,32
СЛ	3,78	3,57	4,16	3,81	3,72	3,46

Таблица 2

Коэффициент	Номер сообщества				
	4	9	13	16	18
ЛР2	0,16	0,17	0,15	0,10	0,20
КД2	1,26	1,18	1,41	1,48	1,32
КОД	0,34	0,33	0,37	0,40	0,44
ЛС1	2,28	2,65	2,35	2,54	2,06
СЛ	3,39	3,12	3,41	3,38	2,90

Заключение

Выделение неявных сообществ с помощью метода Галактик позволяет получить пересекающиеся сообщества, что крайне актуально для сети Telegram-каналов, многие из которых имеют разную тематическую направленность. Это качество такой сети указывает на необходимость выделения для дальнейшего анализа именно пересекающихся сообществ.

К объединенным текстам из Telegram-каналов выделенных сообществ применен анализ психолингвистических факторов для оценки качества этого разбиения на сообщества. Результаты исследования показывают, что значения психолингвистических характеристик текстов отличаются для различных выделенных неявных сообществ, даже для сообществ близкой тематики.

Таким образом, показана возможность оценки корректности выделения неявных сообществ с помощью методов анализа текстов и определения их психолингвистических характеристик средствами компьютерной лингвистики.

Данная механика показывает возможности для анализа информационного воздействия при сетевых взаимодействиях в сети Telegram-каналов.

Список литературы

1. Yarushkina N., Filippov A., Moshkin V., Namestnikov A., Guskov G. The social portrait building of a social network user based on semi-structured data analysis. CEUR Workshop Proceedings/ 14th International Conference on Interactive Systems: Problems of Human-Computer Interaction, 2019. Vol.2475, 2019, P.119-129.
2. Принятие решений на основе анализа слабоструктурированных данных социальных сетей / Н.Г. Ярушкина, И.А. Андреев, А.С. Желепов, В.С. Мошкин – Ульяновск : УлГТУ, 2022. – 113 с.
3. Fortunato S., Newman M.E.J. 20 years of network community detection. Nat. Phys. 2022. 18. P. 848–850.
4. Лещёв Д.А., Сучков Д.В., Хайкова С.П., Чеповский А.А. Алгоритмы выделения групп общения // Вопросы кибербезопасности. 2019. Т. 32. № 4. С. 61-71.
5. Соколова Т.В., Чеповский А.А. Анализ профилей сообществ социальных сетей // Системы высокой доступности. 2018. Т. 14. № 3. С. 82-86.
6. Коломейченко М.И., Поляков И.В., Чеповский А.А., Чеповский А.М. Выделение сообществ в графе взаимодействующих объектов // Фундаментальная и прикладная математика. 2016. Том 21. №3. С. 131-139.
7. Roth M., Ben-David A., Deutscher D. Suggesting Friends Using the Implicit Social Graph — KDD'10. July. 25–28. 2010. Washington. DC. USA. 2010.
8. Girvan M., Newman M. Community structure in social and biological networks // Proceedings of the National Academy of Sciences. 2002. Vol. 99. No. 12. P. 7821-7826.
9. Blondel V.D., Guillaume J.L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks // Journal of Statistical Mechanics: Theory and Experiment. 2008. No. 10. P 10008.
10. Rosvall M. The map equation / M. Rosvall, D. Axelsson, C. T. Bergstrom // The European Physical Journal Special Topics. 2009.
11. Chepovskiy A.A., Leshchev D.A., Khaykova S.P. Core Method for Community Detection / Complex Networks & Their Applications IX. Volume 1: Proceedings of the Ninth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2020. Springer. 2021. P. 38-50.
12. Blöcker, C., Nieves, J.C. & Rosvall, M. Map equation centrality: community-aware centrality based on the map equation. Appl Netw Sci 7, 56 (2022).
13. Rajeh, S., Savonnet, M., Leclercq, E. et al. Comparative evaluation of community-aware centrality measures. Qual Quant 57, 1273–1302 (2023).
14. Попов В.А., Чеповский А.А. Модели импорта данных из мессенджера Telegram // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2022. Т.20. №2. С. 60-71.

15. Попов В.А., Чеповский А.А. Выделение неявных пересекающихся сообществ на графе взаимодействия Telegram-каналов с помощью «метода Галлактик» // Труды ИСА РАН. Том 72. 4/2022. С. 39-50.
16. Аванесян Н.Л., Соловьев Ф.Н., Чеповский А.А. Характеристики текстов сообществ социальных сетей // Вестник НГУ. Серия: Информационные технологии. 2021 Т.19, №1. С.5-14.
17. Чеповский А.М. Информационные модели в задачах обработки текстов на естественных языках. – М.: «ИНТУИТ», 2015. – 228 с.
18. Чеповский А. М. Методы компьютерной лингвистики и анализ текстов на естественных языках. Учебное пособие. М.: «ИНТУИТ», 2015. – 276 с.
19. Смирнов И.В., Шелманов А.О., Кузнецова Е.С., Храмоин И.В. Семантико-синтаксический анализ естественных языков. Часть II. Метод семантико-синтаксического анализа текстов // Искусственный интеллект и принятие решений. М.: ИСА РАН - 2014. - №1 - С. 11-24.
20. Соловьев Ф. Н. Автоматическая обработка текстов на основе платформы ТХМ с учетом анализа структурных единиц текста // Вестник НГУ. Серия: Информационные технологии. 2020. Т. 18, № 1. С. 74-82. DOI 10.25205/1818-7900-2020-18-1-74-82
21. Ениколопов С. Н., Кузнецова Ю. М., Смирнов И. В., Станкевич М. А., Чудова Н. В. Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований. Часть 1. Методические и методологические аспекты // Искусственный интеллект и принятие решений. – 2019. – №. 2. – С. 28-38.
22. Egorova E.E, Chepovskiy A.M., Lavrentiev A.M. A structural pattern based method for automated morphological analysis of word forms in a natural language. // Journal of Mathematical Sciences. 2016. Vol. 214. No. 6. P. 802-813.

ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В ЗАДАЧЕ ВЫЯВЛЕНИЯ МОШЕННИЧЕСКИХ ДЕЙСТВИЙ В ОТНОШЕНИИ КЛИЕНТОВ БАНКА ПРИ ПОДТВЕРЖДЕНИИ ОПЕРАЦИИ

Шелепова А.Н., Воробьев И.А.
Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

На сегодняшний день выявление мошенничества в банковской сфере значительно затруднено из-за применения злоумышленниками методов социальной инженерии. Мошенники обманывают клиентов и убеждают перевести денежные средства на свои счета под различными предлогами. В целях противодействия угрозе банки блокируют операции и обращаются к клиенту для дополнительного подтверждения.

Находясь под психологическим воздействием злоумышленников, клиенты подтверждают операции, несмотря на предупреждения от банка о выявленном мошенничестве. Для поиска подобных сценариев мошенничества в данном исследовании используется метод «дерево решений» (Decision Tree) с применением техники SMOTE для нивелирования дисбаланса классов при обучении модели. Предлагаемый подход повышает качество системы противодействия мошенничеству и позволяет обрабатывать подозрительные операции более эффективно в случаях, когда клиенты банков находятся под сильным психологическим воздействием.

Введение

С каждым годом растет число пользователей онлайн-банкинга. Это связано с увеличением числа пользователей смартфонов и с развитием банковских технологий. Популярность использования онлайн-банкинга клиентами привела к расширению возможностей при совершении киберпреступлений и применению в схемах новых способов обмана. Каждый год мы наблюдаем рост мошеннических действий [1]. Это ведет к финансовым потерям у банков и их клиентов.

Большинство финансовых институтов разрабатывают автоматизированные системы для анализа транзакций, так называемые системы фрод-мониторинга. Их основное назначение – выявление противоправных действий и оповещение о мошенничестве клиентов. Особенностью мошенничества является его адаптируемость под новые условия, выстраиваемые банками для противодействия. Это приводит к частым изменениям паттернов мошенничества, что влияет на качество классификации модели фрод-мониторинга и при отсутствии реакции банка основные характеристики выявления мошенничества деградируют. Адаптируемость заключается и в «психологическом» обходе систем фрод-мониторинга, проявляющегося в обмане клиентов. Мошенники, используя методы социальной инженерии, убеждают клиента подтвердить операцию как легитимную после запроса банка о подозрительности операции. Клиент сообщает о мошенничестве лишь спустя время после подтверждения операции, уже после того, как система фрод-мониторинга разрешила такую операцию. У банков не всегда есть достаточные финансовые и человеческие ресурсы, чтобы напрямую связаться с каждым клиентом через оператора колл-центра и «противостоять» убедительности мошенника. Большинство подтверждений проводятся через приложения или роботизированный звонок и являются достаточно номинальными и неэффективными в том случае, если клиент уже убежден мошенником. Решение задачи выявления мошенничества в подтвержденных операциях может помочь в определении клиентов, для которых требуется более индивидуальный подход при обнаружении мошенничества и необходимо подключение обученного оператора в процесс разубеждения клиента.

Построение модели и оценка качества классификации

Для оценки эффективности предлагаемого подхода выбраны традиционно используемые в задачах выявления мошенничества метрики.

$$Precision = \frac{TP}{TP + FP'}$$

$$Recall = \frac{TP}{TP + FN'}$$

где TP (True positive) – мошенническая операция, идентифицированная корректно как мошенническая; FP (False positive) – легитимная операция, идентифицированная как мошенническая; TN (True negative) – легитимная операция, идентифицированная корректно; FN (False negative) – мошенническая операция, идентифицированная как легитимная.

Recall (полнота) позволит оценить долю мошенничества, выявленную классификатором по отношению ко всем мошенническим операциям; Precision (точность) – вероятность того, что подозреваемая классификатором операция действительно мошенническая.

Обычно увеличение полноты ведет к уменьшению точности. Необходимо находить баланс между двумя метриками, в том числе, исходя из особенностей реальной задачи. Метрика, позволяющая оценить одновременный рост точности и полноты, является *F*-мера (*F*), которая представляет собой взвешенное гармоническое среднее двух характеристик:

$$F = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Другой способ увидеть взаимоотношение точности и полноты – это построение PR-кривой (кривая точности-полноты). Численную оценку качества модели можно получить через площадь под данной кривой (PR-AUC). Чем больше площадь под кривой, тем в среднем лучше классификатор.

В работе используется алгоритм «дерево решений» (*Decision Tree*) – алгоритм обучения с учителем для классификации и регрессии, преимуществом которого для банковской сферы является интерпретируемость [2]. Для поиска оптимальных значений параметров модели используется встроенный метод решетчатого поиска с перекрёстной проверкой GridSearchCV из библиотеки *scikit-learn* [3]. Для борьбы с дисбалансом данных применяется реализация техники генерации искусственных примеров SMOTE [4] из библиотеки *imblearn*.

В работе применяется out-of-time тестирование, то есть тестирование на данных вне времени обучения. Это необходимо для того, чтобы понять, насколько предсказательная способность модели ухудшается со временем, так как она может терять свою разделительную способность с течением времени из-за изменения паттернов мошенничества.

Эксперименты на реальных данных и результаты

Описание данные. Данные предоставлены крупной финансовой организацией в обезличенном формате – по ним невозможно получить какую-либо информацию о клиенте, с которым связана транзакция. Также названия признаков замаскированы и не указывают на происхождение данных. Массив данных содержит 277,208 примеров и 221 признак. Данные представляют собой набор транзакций, где каждая транзакция была оценена существующей системой фрод-мониторинга. В исследуемый набор данных в качестве признаков для моделирования вошли условия, использующиеся для настройки системы правил принятия решения по операции на предмет мошенничества. В случае выполнения условия соответствующему признаку присваивалось значение 1, если не выполняется, то 0. Дополнительно система фрод-мониторинга фиксировала дату, свое решение о том, блокировать (**deny**) или пропустить (**allow**) транзакцию, итоговую резолюцию о том, является ли транзакция мошеннической (**F**) или легитимной (**G**). Также в массиве данных зафиксировано, какая реакция фрод-мониторинга на транзакцию была выбрана. Реакции соответствуют следующим моделям поведения:

- роботизированный звонок клиенту, если система фрод-мониторинга приняла решение, что операция подозрительная;
- звонок оператора, если система фрод-мониторинга приняла решение, что операция подозрительная;
- другие способы связи (например, SMS, подтверждение транзакции в мобильном приложении и т.д.);
- обеляющее правило, означающее, что операция была подтверждена клиентом после роботизированного звонка;
- обеляющее правило, означающее, что операция была подтверждена после звонка оператора.

Классы. Для выявления «разрешенных» системой фрод-мониторинга операций, которые затем оказались мошенническими, положительным классом были размечены транзакции с решением о проведении **allow** и резолюцией **F**. В отрицательный класс вошли следующие транзакции: отклоненная системой операция, которая оказалась мошеннической; отклоненная операция, которая оказалась легитимной; разрешенная легитимная.

Подтверждение актуальности. После разметки данных количество «неверно» разрешенных к проведению фрод-мониторингом операций при звонке робота составило 7,876, при звонке оператора клиенту 118. Такое соотношение подтвердило, что мошенники пытаются построить сценарии обмана клиента в автоматизированные проверки банка.

Выборки. Для обучения использовались наборы данных X_{train} , y_{train} , включающие 80% транзакций, произошедших за календарную неделю. Оставшиеся 20% транзакций были выделены для тестирования – выборки X_{test} , y_{test} . Дополнительно, для проверки устойчивости предложенного подхода во

времени рассмотрены операции, произошедшие за календарную неделю после недели, на которой проводится обучение. Таким образом получены данные для out-of-time тестирования – это выборки $X_{test_out_of_time}$, $y_{test_out_of_time}$.

Баланс классов. Из выбранных примеров для обучения 86,1% (21,203) являлись отрицательным классом и только 13,8% (3,400) транзакций в наборе данных являлись мошенническими, что указывает на дисбаланс классов. Для недели, на которой осуществлялась out-of-time проверка соотношение составило 89,3% отрицательного класса к 10,6% положительному.

Выбор признаков. Был отобран 121 признак следующим образом – его среднее значение, подсчитанное по всему набору данных, находилось в диапазоне от 0.02 до 0.984 (граничные значения выбраны экспертно). После такой фильтрации были исключены признаки, которые заполнены полностью нулями или полностью единицами. А также отобраны только признаки, принимающие одно из двух значений, исключая признаки с вещественными значениями. Среди отобранных признаков не было пропущенных значений.

Лучшие модели. В результате были получены две модели со следующими метриками. Первая модель – дерево решений, обученное на несбалансированных данных X_{train} (обозначается DT) с точность – 0.8009 на данных X_{test} и с ухудшением точности до 0.6980 на out-of-time выборке. С помощью решетчатого поиска с перекрестной проверкой были получены следующие параметры:

{'max_depth': 15, 'min_impurity_decrease': 1e-05, 'min_samples_leaf': 16}.

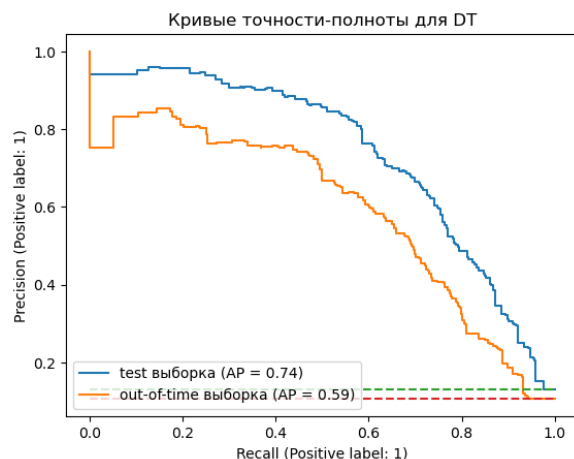


Рис. 1. Сопоставление кривых точности-полноты DT при тестировании на X_{test} и данных через неделю – $X_{test_out_of_time}$

Таблица 1. Метрики для модели DT

Выборка	Точность	Полнота	F	PR-auc
X_{test} , y_{test}	0.8009	0.5862	0.6769	0.7359
$X_{test_out_of_time}$, $Y_{test_out_of_time}$	0.6980	0.5002	0.5828	0.5909

Вторая модель, обученная на данных с добавлением сгенерированных синтетических примеров с помощью техники SMOTE (обозначается DT SMOTE), позволила добиться лучшей полноты – 0.7492 на тестовых данных и 0.6656 на данных out-of-time. В результате были получены следующие параметры модели:

{'criterion': 'entropy', 'max_leaf_nodes': 215, 'min_samples_leaf': 36, 'min_samples_split': 20}.

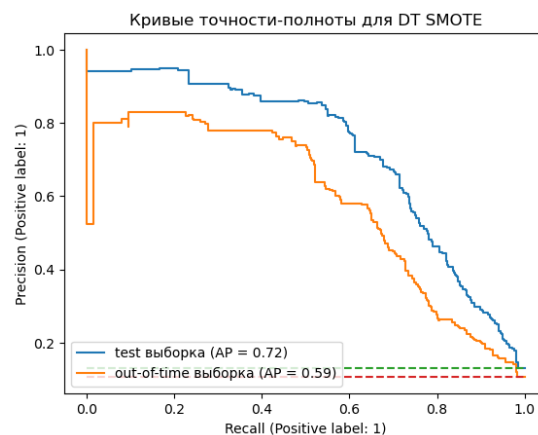


Рис. 2. Сопоставление кривых точности-полноты DT SMOTE при тестировании на X_{test} и данных через неделю – $X_{test_out_of_time}$

Таблица 2. Метрики для модели DT SMOTE

Выборка	Точность	Полнота	F	PR-auc
X_{test} , y_{test}	0.5591	0.7492	0.6403	0.7136
$X_{test_out_of_time}$, $Y_{test_out_of_time}$	0.4529	0.6656	0.5390	0.5815

В результате анализа моделей было получено пересечение множеств их лучших признаков. После экспертного анализа установлено, что их фактический смысл соответствует тому, что клиент старше 70 лет, осуществляется операция на высокую сумму и клиентом взят кредит в день операции. Это отражает текущие тренды мошенников – обманывать пенсионеров, которые более доверчивы, с использованием заемных средств клиентов.

Применимость. Полученные модели легко преобразовываются в набор правил, благодаря интерпретируемости алгоритма и могут быть интегрированы в существующую систему фрод-мониторинга.

Заключение

В данной работе показана эффективность алгоритма дерева решений для выявления мошеннических действий в отношении клиентов банка при подтверждении операции. Получен инструмент, позволяющий выявлять мошенничество, в котором клиент находится под сильным психологическим воздействием, с эффективностью сопоставимой с показателями существующей системы фрод-мониторинга. Используя данный инструмент, финансовые инсти-

туты смогут выявлять более сложные случаи мошенничества и запускать сценарии их обработки, в которых клиенты будут более защищены от психологического воздействия мошенников.

Список литературы

1. Ларионова С. Л. — Организация работы по борьбе с мошенничеством в кредитной организации// Вопросы безопасности. — 2022. — № 4. — С. 15 – 26. DOI: 10.25136/2409-7543.2022.4.38920 EDN: CLHLDR URL: https://nbpublish.com/library_read_article.php?id=38920 (дата обращения: 31.12.22)
2. John F. Magee, Decision Trees for Decision-Making// Harvard. Business Review, July-August, 1964. P. 127.
3. scikit-learn: Machine Learning Python library [Электронный ресурс]. URL: <https://scikit-learn.org/stable/index.html> (дата обращения: 22.01.2023)
4. Chawla N., Bowyer K., Hall L., Kegelmeyer W. SMOTE: Synthetic Minority Over-sampling Technique // Journal of Artificial Intelligence Research, vol. 16, 2002. P. 321-357.

ИССЛЕДОВАНИЕ ЭНТРОПИИ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ МЕТОДОМ СОСТАВЛЕНИЯ СЛОВАРЕЙ N-ГРАММ¹

Малашина А.Г.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

Данная работа является продолжением работ автора [1, 2] по исследованию энтропии n-грамм.

В статье исследуется словарный метод определения энтропии n-грамм и его чувствительность к разным языкам и стилям текстов. Для исследования выбраны русский и английский языки, а также публицистический и художественный стили. Приводятся соответствующие экспериментальные оценки энтропии n-грамм.

Введение

Естественные языки – это языки, которые люди используют для общения. В отличие от искусственных языков они не создавались целенаправленно. Естественный язык, как сложную знаковую систему, характеризует информационная энтропия – степень неопределенности следующего высказывания.

Тексты на естественном языке имеют энтропию значительно меньшую, чем максимальная энтропия его алфавита, в то время как случайный набор символов (случайная n-грамма) обладает максимально

возможной энтропией в данном алфавите. Таким образом, значения энтропии могут быть использованы для автоматического распознавания осмысленного² текста среди случайного набора символов [3].

Существуют различные методы определения энтропии текстов. Наиболее популярным из них является метод Шеннона [4]. Представляя текст цепью Маркова n-го порядка, можно приблизительно оценить вероятности n-грамм. В этой статье предлагается использовать метод на основе словарей n-грамм для определения энтропии, идеи которого восходят к комбинаторному подходу Колмогорова [5].

Известно, что энтропии разных языков различаются. Более того, значение энтропии зависит и от стиля текста. Цель исследования – оценить энтропию n-грамм различных текстов на русском и английском языках и исследовать чувствительность словарного метода к языковым и стилистическим особенностям текстов.

Корпуса текстов

Для исследования энтропии n-грамм составляются языковые корпуса общим объемом 10 млн символов каждый. Для русского языка используются тексты из *Открытого корпуса* [6], для английского – из *iWeb corpus* [7] преимущественно в публицистическом и разговорном стилях.

Чтобы увеличить покрытие словарей n-грамм, размер алфавита ограничивается. Таким образом, создаваемый корпус проходит процесс нормализации.

Нормализация состоит из следующих этапов:

- 1) удаление HTML-тегов;
- 2) приведение в нижний регистр;
- 3) фильтрация текста (удаление всех символов, кроме «а-я» за исключением «ё» (для русского языка) или «a-z» (для английского языка), «.», «,»);
- 4) удаление двойных пробелов, повторяющихся точек и запятых, а также пробелов перед точками и запятыми.

Таким образом, мощность алфавитов корпусов для русского языка составляет 35 символов, для английского – 29 символов. Данные алфавиты – простые расширения кириллического и латинского алфавитов соответственно.

Актуальность исследования связана с важной для криптографических приложений задачей восстановления исходных сообщений по информации о преобразованном сообщении. В данном аспекте значительную роль играет структура источника сообщений и, в частности, количество сообщений заданной длины, которое может быть сгенерировано. Поэтому в данной работе излагается ряд результатов по исследованию структуры отдельных отрезков текста на русском и английском языках ограниченной длины.

Словарный метод определения энтропии

Словарь представляют собой набор n-грамм, расположенных в алфавитном порядке без повторов. N-граммы выбираются из текста с зацеплением: для

¹ Набор из n знаков алфавита языка.

² Осмысленный текст – текст, вырабатываемый естественным источником

следующей n -граммы осуществляется сдвиг вправо на один символ.

Процесс создания словаря включает извлечение всех n -грамм из корпуса, удаление повторяющихся n -грамм и сортировку. Объем словаря — это количество уникальных n -грамм, которые остаются после удаления дубликатов.

В этой статье рассматриваются n -граммы четырех порядков: 10-граммы, 15-граммы, 20-граммы и 25-граммы. Шаг в 5 символов между длинами n -грамм выбран для более точного построения последующей модели экстраполяции.

Для низких порядков n -грамм уже известны некоторые значения энтропии, в отличие от n -грамм порядка свыше 10. При этом провести расчеты для n -грамм порядком выше 25-30 является трудоемкой задачей.

Составленные словари n -грамм составляют основу методологии расчета значений энтропии. Предполагается, что созданные словари являются инструментом для автоматического различения осмысленных n -грамм, существующих в языке, и случайных n -грамм.

Экспериментальная оценка словаря рассматривается как некоторая аппроксимация числа всех возможных n -грамм в языке соответствующего порядка. Это означает, что все n -граммы, отсутствующие в словаре, рассматриваются как случайные тексты. Основываясь на этом предположении, можно приблизительно оценить энтропию n -грамм.

В рамках n -граммной модели языка текст представляет собой реализацию независимых испытаний, результатами которых являются n -граммы соответствующего естественного языка. Энтропия случайного источника может быть вычислена методом Шеннона [4]. Однако, чтобы избежать вычисления вероятностей n -грамм, предлагается комбинаторный подход к вычислению энтропии на основе объема словаря. Идея метода восходит к комбинаторному методу Колмогорова [5] и второй теореме Шеннона.

Пусть объем K_n словаря n -грамм является приближением числа всех возможных n -грамм в языке. Тогда удельная энтропия n -грамм на символ (бит/символ) может быть оценена как: $H_n = \frac{\log_2 K_n}{n}$.

Предел данной величины является приближением энтропии соответствующего языка.

Энтропия русского и английского языков

На диаграмме (рис. 1) представлены результаты оценки словарей n -грамм для русского и английского языков. Объемы словарей русского языка превышают соответствующие значения для английского, однако эта разница наиболее заметна при меньшей длине n -грамм. С увеличением n разница в объемах словарей разных языков начинает сглаживаться.

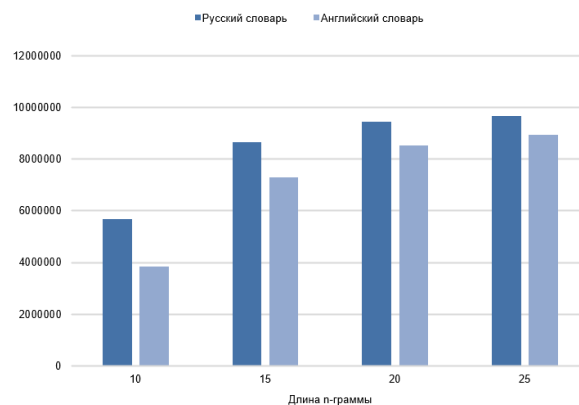


Рис. 1. Объем словарей n -грамм

На основе данных об объеме словарей, рассчитана энтропия коротких n -грамм. Объем информации, передаваемой одной n -граммой, увеличивается с увеличением ее длины. На графике (рис. 2) можно видеть, что удельная энтропия текста уменьшается с увеличением длины n -граммы.

С ростом длины n -грамм значение энтропии уменьшается до некоторого n и при дальнейшем росте почти не меняется, то есть достигает определенного предела, называемого энтропией языка. Экстраполировать эти результаты на большие значения n сложно, поскольку форма этой последовательности значений, как правило, неизвестна, за исключением того, что она положительно уменьшается. Чтобы получить предельную энтропию из этого набора измерений, строится модель последовательных оценок.

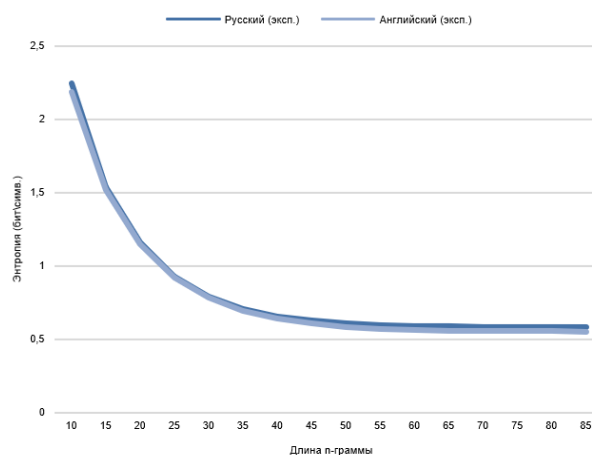


Рис. 2. Энтропия n -грамм русского и английского языков

Предположительно последовательность значений энтропии подчиняется линейному рекуррентному соотношению: $F_n - F_{n+1} = k \cdot (F_{n+1} - F_{n+2})$ с начальными условиями $F_0 = H_{10}, F_1 = H_{15}$ и $F_2 = H_{20}$.

Коэффициент k для модели определяется численно в соответствии с экспериментально полученными значениями энтропии для n -грамм малой длины. Наилучшее приближение достигается при $k \approx 0,6$. Путем увеличения значения n строится последовательность эвристических оценок H_n , эксперимен-

тальная оценка которых становится затруднительной при большой длине n -граммы. Начиная с $n = 50$, значения H_n стабилизируются и больше не изменяются с ростом длины n -граммы.

Таким образом, определяются энтропии соответствующих языков (табл. 1).

Таблица 3. Значения энтропии языков.

Язык	Энтропия (бит/симв.)
Русский язык	0,61
Английский язык	0,59

Энтропия английского языка несколько ниже, чем соответствующее значение для русского. Однако такая разница для языков совсем незначительна, что говорит о том, что словарный метод *слабо чувствителен* к разным языкам. Кроме того, полученные результаты ниже реальных значений энтропии, поэтому их можно рассматривать как нижнюю оценку энтропии языка.

Использование марковских источников для вычисления значений энтропии позволяет получить более точные результаты по сравнению со словарным методом [8].

Энтропия текстов разных стилей

Для исследования энтропии текстов разных стилей используется дополнительный корпус художественных текстов, составленный из русских произведений XIX-XX вв. [8]

На диаграмме (рис. 3) представлено сравнение соответствующих значений энтропии n -грамм для публицистических и художественных текстов. Значения практически не различаются, что говорит о том, что словарный метод определения энтропии *не чувствителен* к изменению стиля текста.

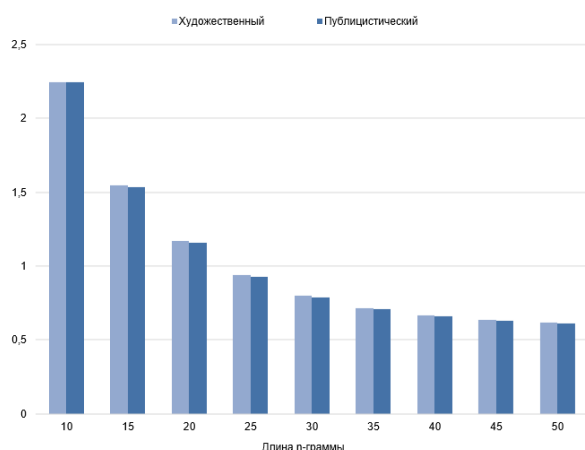


Рис. 3. Энтропия n -грамм текстов разных стилей

Заключение

В работе был рассмотрен словарный метод определения энтропии n -грамм и соответствующих языков. Для исследования метода было проведено сравнение значений энтропии русского и английского языков. Было установлено, что словарный метод практически не чувствителен к разнице в энтропии

языков и различных стилей. С помощью полученных значений можно приблизительно оценить нижнюю границу соответствующих языков. Для получения более точных значений энтропии стоит использовать метод на основе марковских цепей. Однако применение словарного метода может быть целесообразно для задач, где не требуется высокая точность оценки энтропии, например для проведения атак по словарю.

Список литературы

1. Малашина А. Г., Лось А. Б. Построение и анализ моделей русского языка в связи с исследованиями криптографических алгоритмов // Чебышевский сборник. – 2022. Т. 23. № 2. – С. 151-160.
2. Malashina A. The Combinatorial Analysis of n -Gram Dictionaries, Coverage and Information Entropy based on the Web Corpus of English // Baltic Journal of Modern Computing. – 2021. Vol. 9. No. 3. – P. 363-376.
3. Яглом А. М., Яглом И. М. Вероятность и информация. // Издательство «Наука». – 1973.
4. Shannon C. E. A mathematical theory of communication // The Bell system technical journal. – 1948. – №27(3). – P. 379–423.
5. Колмогоров А. Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации– 1965. – Т.1. №1. – С. 3-11.
6. OpenCorpora: открытый корпус русского языка. Источник: <http://opencorpora.org>.
7. iWeb - English-Corpora. Источник: <https://www.english-corpora.org/iweb/>.
8. Нагаева И. Э., Савченкова Д. М. Исследование информационных характеристик художественных текстов и их переводов // Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов имени Е.В. Арменского. Материалы конференции. – Москва, МИЭМ. – 2021. – С. 241-244.

ПРИМЕНЕНИЕ СОСТОЯТЕЛЬНЫХ МЕР ЗАВИСИМОСТИ К РЕШЕНИЮ ЗАДАЧИ ОБНАРУЖЕНИЯ ИНСАЙДЕРА

Эшкинин А.И.¹, Доманова В.С.²,
Фонарева О.Е.², Мартынов А.С.²

Национальный исследовательский университет
«Высшая школа экономики»,
¹ департамент прикладной математики,
² департамент электронной инженерии
МИЭМ НИУ ВШЭ

Аннотация

В работе рассматривается новый подход выявления внутренних злоумышленников в информационной системе, связанный с использованием энтропийных мер. Выявление различий между поведением внутреннего злоумышленника и легитимного сотрудника и оценивание данных различий с помощью статистических мер может позволить на ранних этапах выявить потенциальную угрозу.

Введение

Одной из важных задач обеспечения информационной безопасности является проблема намеренных или случайных инсайдерских атак. Сложность их обнаружения заключается в том, что внутренний злоумышленник уже имеет доступ к информационной системе, что делает большое количество средств безопасности бессильными, к тому же нелегитимные действия работника часто происходят в рамках данных ему полномочий. Несмотря на существование современных систем предотвращения утечек данных, использующих морфологический анализ, цифровые отпечатки и прочие методы, количество инцидентов утечки данных, вызванных действиями инсайдеров стремительно растет, так же, как и ущерб от них.

Изучаемый подход основан на статистическом оценивании различий между действиями сотрудника и эталонными.

Анализ проблемной области

Для борьбы с угрозой внутреннего нарушителя было проведено много исследований. В США на государственном уровне в начале прошлого десятилетия была создана программа Anomaly Detection at Multiple Scales (ADAMS) [1], целью которой являлось создание и применение технологии для решения проблемы обнаружения аномалий в больших наборах данных. Алгоритм работы программы включал четыре основных этапа:

1. Обнаружение аномалий на основе структурированных особенностей данных. Это позволяет значительно сократить объем данных, но по-прежнему может привести к слишком большому количеству ложных срабатываний, с которыми операторы не смогут справиться.

2. Применение семантических данных. Программное обеспечение использует вспомогательную информацию из различных источников, чтобы сформировать семантический фон, на основе которого будут интерпретироваться данные. Затем семантический слой применяется к обнаруженным структурным аномалиям, чтобы снизить частоту ложноположительных результатов до уровней, управляемых человеком-оператором.

3. Ранжирование угроз. Программное обеспечение будет принимать результирующее аномальное поведение и ранжировать его по важности (в интересах человека-оператора).

4. Получение обратной связи от операторов. Операторы будут помечать каждую угрозу как реальную атаку или как ложное срабатывание. На основе этих данных программное обеспечение будет обучаться.

Основной недостаток данной системы заключается в том, что она основана на технологии машинного обучения. Для обучения программы требуется хранить огромные объемы данных.

В современном мире в компаниях для защиты от внутреннего злоумышленника используются DLP-системы, работающие по следующему принципу: система определяет, где находятся конфиденциальные данные, которые требуется защищать; произво-

дится мониторинг выявленных данных, то есть отслеживается, как данные используются; при нарушении заранее установленных правил нелегитимное действие либо останавливается, либо информация о действии записывается для анализа инцидента уже после его совершения.

Несмотря на такой подход к защите, внутренний злоумышленник может попытаться обойти DLP-систему. Если система настроена с ошибками, то злоумышленник может этим воспользоваться при совершении нелегитимных действий. Это могут быть простые способы эксфильтрации данных с обходом DLP, например, отправка файлов на личную почту, конфиденциальная информация в которых была видоизменена так, чтобы DLP-система не смогла распознать её. Более осведомленные злоумышленники могут использовать более сложные атаки, содержащие в себе большое количество тактик, позволяющих украсть защищаемые данные.

В своей работе мы предлагаем анализировать действия пользователей в поточном режиме с помощью статистических мер зависимости, таких как энтропия Шеннона.

Похожие исследования уже проводились в научном сообществе. Например, в диссертации [2] рассматривается возможность выявления инсайдера статистическими методами. Для этого информационная система компании рассматривается как хранилище данных. Работой сотрудников в компании являются запросы данных из хранилища. Нелегитимным действием является запрос данных, которые не требуются для выполнения функциональных обязанностей сотрудника, причем злоумышленник нацелен на данные, к которым у него нет доступа. Такая модель исключает ошибки первого рода, а выявление потенциального инсайдера будет выполнено за конечное время. Недостаток данной системы заключается в том, что необходимо всю информационную систему компании представить в виде хранилища данных и распределить, доступ к каким аспектам информации может быть предоставлен каждому сотруднику.

Практическое исследование

В статье “Bridging the Gap: A Pragmatic Approach to Generating Insider Threat Data” авторы рассуждают о сборе данных для анализа [3]. Чтобы быть полезными, эти данные должны содержать подробный отчет о поведении человека в контролируемой среде. В конце концов, внутренние угрозы связаны с действиями людей, а не машин, и методы обнаружения неизбежно будут включать методы социальных наук. Но такие наборы данных трудно получить. У исследователей в этой области есть два варианта - они могут использовать или собирать реальные пользовательские данные, или они могут использовать синтетические данные. Поскольку злонамеренные инсайдеры — это, прежде всего, инсайдеры, для сбора реальных данных некоторая организация должна непосредственно отслеживать и записывать поведение и действия своих собственных сотрудников. Конфиденциальность и соображения неприкосно-

венности частной жизни создают барьеры для сбора и использования таких данных в исследовательских целях. Таким образом, иногда предпочтительнее использовать синтетические данные. Благодаря развитой системе генерации синтетических данных, подобной той, что используется здесь, пользователь может гибко контролировать и быстро и экономично генерировать наборы данных с желаемыми характеристиками, размером и качеством относительно измеримых характеристик.

Для исследования был взят публичный набор искусственно созданных данных “Insider Threat Test Dataset” от университета Карнеги-Меллон [4]. В выборке представлены действия 1000 пользователей, из которых 70 инсайдеров, действующих по 1 из 5 заранее прописанных сценариев.

Исследование проводилось по 5 параметрам: вход и выход из системы (854859 записей), подключение и отключение съемных носителей (405380 записей), перемещение файлов на съемные носители (445581 записей), отправка электронных писем (2629978 записей) и посещение интернет-сайтов (854859 записей).

Для оценивания различий между поведением инсайдера и легитимного сотрудника предлагается использование дивергенции Кульбака-Лейблера:

$$D_{KL}(P||Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i},$$

где P - априорное распределение, Q - апостериорное распределение. Данная мера позволяет оценить, насколько информационная энтропия одного распределения отличается от другого. В качестве априорного распределения используется распределение вероятностей, полученных на основе усредненных данных по всем легитимным сотрудникам. Апостериорное распределение - распределение сотрудника, который проверяется на нелегитимные действия.

Производится расчет дивергенции для каждого параметра. Вероятность случайного события для посещения интернет-сайтов рассчитывается как отношение количества посещений определенного сайта и количества посещений всех сайтов за период времени. Вероятность для других параметров рассчитывается как количество таких действий пользователем за определенный временной период, деленное на количество действий пользователем за весь период.

На нижеприведенных графиках каждый столбец диаграммы соответствует значению дивергенции между действиями некоторого пользователя и усредненным легитимным набором данных. Красным цветом отмечены инсайдеры, синим - легитимные пользователи.

На рисунке 1 представлен график подключений и отключений сторонних девайсов к устройствам пользователей. 90%-квантиль для правомерного пользователя составляет 0.18, для инсайдера такое значение имеет 10%-квантиль.

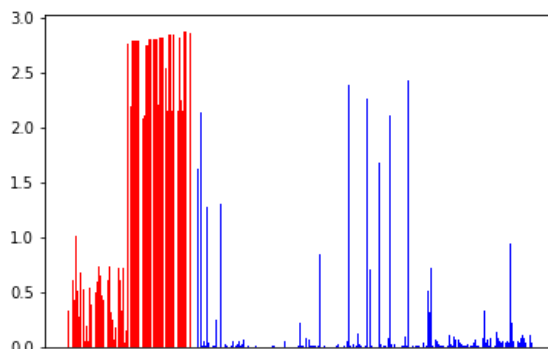


Рис. 1. График значений дивергенции для событий подключения сторонних девайсов сотрудниками

Значения дивергенций для действий по копированию файлов на съемные носители показаны на рисунке 2. 88%-квантиль легитимного пользователя, равен 10%-квантилю инсайдера и составляет 0.20.

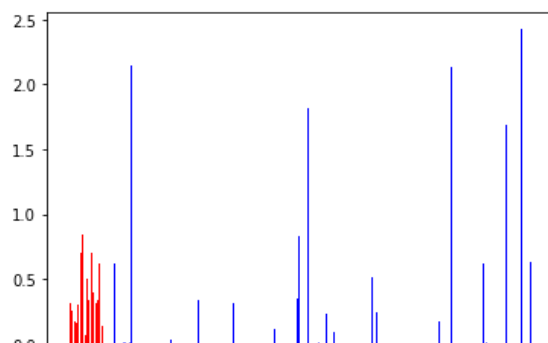


Рис. 2. График значений дивергенции для событий по перемещению файлов

На рисунке 3 представлен график входов и выходов из системы. 92%-квантиль легитимного пользователя, равен 9%-квантилю инсайдера и составляет 0.1.

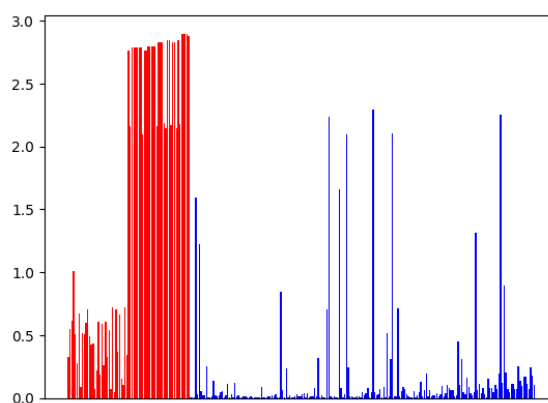


Рис. 3. График значений дивергенции для событий входа/выхода из системы (красным цветом отмечены инсайдеры, синим - легитимные пользователи)

На рисунке 4 представлен график интернет-трафика. 90%-квантиль для легитимного пользователя составляет 6.81, 10%-квантиль для инсайдера 4.05.

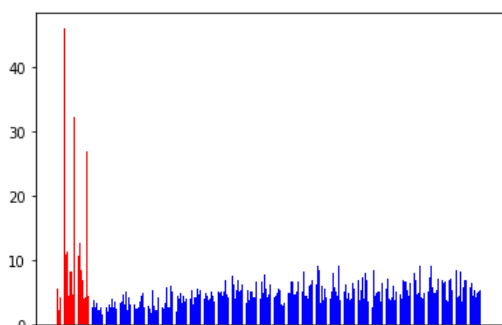


Рис. 4. График значений дивергенции для интернет-трафика

На рисунке 5 представлен график отправки электронных писем. 92% и 10% квантили для правомерного пользователя и инсайдера равны соответственно 0.1 и 0.12.

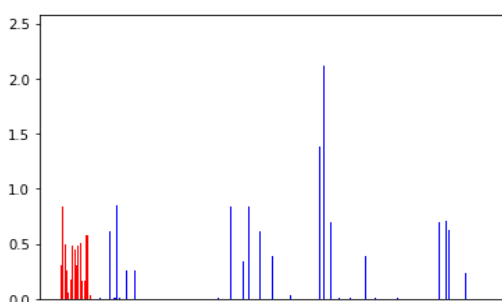


Рис. 5. График значений дивергенции для электронных писем

Заключение

В ходе исследования был рассмотрен метод обнаружения инсайдера, основанный на вычислении значения дивергенции. Полученные результаты показывают значительные различия между действиями инсайдеров и правомерных пользователей. Дивергенция абсолютного большинства инсайдеров (~90%) относительно эталонного сотрудника не опускается ниже значения 0.1-0.2, тогда как для правомерных пользователей такие значения достигаются лишь в ~10% случаях. Данный проект может быть использован для последующих исследований способов обнаружения инсайдера.

Список литературы

1. Anomaly Detection at Multiple Scales (ADAMS) // URL: https://www.govloop.com/wp-content/uploads/2015/07/DARPA-BAA-11-04_ADAMS-2.pdf (дата обращения: 27.12.2022).
2. Мартынов Е.А. Исследование и разработка методик оценки защищенности информационных объектов от потенциальных нарушителей: дис. канд. техн. наук: 05.13.19. - Москва, 2018. - 132 с.
3. Glasser J., Lindauer B. Bridging the gap: A pragmatic approach to generating insider threat data // IEEE Security and Privacy Workshops. - 2013. - №5. - С. 98-104.
4. Lindauer, Brian (2020): Insider Threat Test Dataset. Carnegie Mellon University. Dataset. "Insider Threat Test Dataset," Cmu.edu, Nov. 28, 2016. <https://doi.org/10.1184/R1/12841247.v1>

ИССЛЕДОВАНИЕ АНОМАЛЬНЫХ СИТУАЦИЙ СЕТЕВОГО ТРАФИКА ПО ДАННЫМ ЖУРНАЛОВ WEB-СЕРВЕРА NGINX

Федоров С.А.

Национальный исследовательский университет
«Высшая школа экономики»,
департамент прикладной математики
МИЭМ НИУ ВШЭ

Аннотация

Предлагается алгоритм поиска аномальных сообщений в журналах событий веб-сервера nginx. Для исследования были взяты данные с промышленного веб-сервера nginx за суточный промежуток. Алгоритм основывается на начальном выделении аномальных сообщений на этапе предварительной обработки и алгоритме машинного обучения без учителя «изолированный лес». В результате алгоритм показал хорошую скорость работы с достаточно высоким показателем точности.

Введение

Под аномальными событиями будем понимать такие экземпляры записей в журнале событий, которые не соответствуют стандартному поведению системы.

С каждым годом количество устройств возрастает и размеры журналов событий становятся больше. Современные подходы и продукты для поиска аномалий в журналах событий, которые имеют разрозненную структуру и большой размер, имеют недостатки.

Основными проблемами развития алгоритмов поиска аномалий в журналах событий можно выделить следующие пункты:

- записи журналов не структурированы, так как разработчики записывают данные в удобном для них виде 1;
- разнообразие шаблонов таких записей и появление новых шаблонов в развивающихся системах;
- записи, ориентированы на восприятие специалистом, что мешает машинному обучению и вынуждает заниматься предварительной обработкой данных и приведением типов.

Цель работы заключается в разработке алгоритма выделения аномальных сообщений в журнале событий веб-сервера nginx. Для достижения цели необходимо выделить аномалии на этапе предварительной обработки, выбрать представление данных для анализа поведения пользователей и подобрать модель машинного обучения.

Алгоритм поиска аномалий

Задачу поиска аномалий составим следующим образом:

Пусть задано множество категорий $A = \{-1, 1\}$, где -1 — это аномальное значение, 1 — это нормальное значение. Пусть X — это множество сообщений в журналах событий веб-сервера nginx. Задача работы состоит в классификации элементов выборки X , то есть, в нахождении неизвестной функции $a^*: X \rightarrow A$. Необходимо разработать алгоритм, который спосо-

бен определить категорию каждого элемента выборки $x \in X$.

Для релевантного исследования были получены реальные журналы событий с промышленного веб-сервера nginx. Была написана программа на языке программирования python для предварительной обработки, добавления поля о стране пользователя в журнал событий и перевода в формат csv для удобного дальнейшего анализа.

Приведем пример необработанной записи журнала событий:

```
89.107.124.36 www.suit.ru - [19/Feb/2022:03:41:07
+0300] "GET / HTTP/1.1" 200 158377 "-" "Mozilla/5.0
(Windows NT 10.0; Win64; x64) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/75.0.3770.100 Safari/537.36" "-" ECDHE-
RSA-AES128-GCM-SHA256 0.100
```

Данная запись в виде записи в таблице csv будет иметь следующий вид (рис. 1).

remote_addr	URL	remote_usr	date_time	date	time	request	req_method	req_uri	http_v
89.107.124.36	www.suit.ru	<NA>	2022-02-19 03:41:07+03:00	2022-02-19	03:41:07	"GET / HTTP/1.1"	GET	/	HTTP/1.1
noinfo	status	body_byte_sent	http_referer	user_agent	forward_for	cipher_suite	request_processing_time	remote_addr	http_v
<NA>	200	158377	<NA>	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/75.0.3770.100 Safari/537.36	<NA>	ECDHE-RSA-AES128-GCM-SHA256	0.1		

Рис. 1. Запись в таблице csv

С помощью библиотеки `geoip2` 5 дополним начальные данные о пользователе, а именно добавим столбец с информацией о стране, откуда идет запрос веб-ресурса. Предлагается алгоритм, состоящий из двух этапов: предварительной обработки с начальным выделением явных аномалий и дальнейшем анализом событий для выявления аномалий.

Этап 1. Предварительная обработка и начальное выделение явных аномалий:

Создадим столбец с названием “anomaly”, в который будем сохранять значение аномальности. В предварительную обработку добавим следующие проверки аномальности:

- Если **URL** отличается от рабочего, то сразу ставим значение, что сообщение аномально.
- Если **remote_usr** имеет не пустое значение, а также не принадлежит списку зарегистрированных пользователей, то тоже ставим, что сообщение аномально.
- Если поля **cipher_suite** и **http_ver** оказались пустыми или не принадлежат списку доверенных значений, то возвращаем -1.
- Если в поле **request** при проверке регулярными выражениями на наличие XSS, SQL-injection и других популярных атак, срабатывает хотя бы одно регулярное выражение ставится -1.

В дальнейшем те данные, которые были определены как аномальные ввиду анализа перечисленных полей, не будут участвовать в обучающей выборке. Все строки, в которых получили -1 в столбце anomaly, вырезаем в отдельную таблицу.

Пользователей будем определять по конкатенации полей **ip** и **user_agent** (информация о браузере, с которого заходят на веб-ресурс). Используя информацию поля **user_agent**, можно понять всю вспомогательную информацию о пользователе, такую как: операционная система пользователя, версия веб-агента, которую он использует, является ли этот пользователь автоматизированной системой (например, Zabbix или ботом) и т. п.

Создадим столбец с конкатенацией столбцов **ip** и **user_agent**, назвав его **Пользователь (User)** и удалим столбцы **ip** и **user_agent**. Поле **cipher_suite** — это совокупность алгоритмов, используемых в конкретной TLS-сессии. Данное поле уже учитывалось на этапе предобработки данных, поэтому оно не будет включено в итоговом выборе. Так как столбец **cipher_suite** уже учли, то в дальнейшем не будем рассматривать его в выборке на следующем этапе.

Этап 2. Выделение аномалий

На этом этапе создается еще одна подобная таблица по пользователям за промежутки времени по 30 минут, состоящую только из количества запросов за данный промежуток времени, суммарное количества байт, переданных за 30 минут и суммарное количество времени, потраченное сервером на обработку каждого запроса пользователя. Нормируем все данные, поделив каждый элемент на сумму всех значений, соответствующего столбца (таблица 1).

Таблица 1. Информация по пользователям веб-сервера за 30 минут в процентах

Пользователь	Количество запросов за промежуток	Суммарное количества байт	Суммарное время ответа в секундах
Пользователь_1	2,74%	1,929%	0,12%
Пользователь_2	0,022%	0,0385%	0,16%
...
Пользователь_n	0,438%	0,0192%	0,008%
Суммарное кол-во	100%	100%	100%

Создадим таблицу за сутки, объединив таблицу 1 за 48 временных промежутков.

Для решения поставленной задачи будем использовать алгоритм машинного обучения изолированный лес. Изолированный лес 2 — это алгоритм машинного обучения без учителя, один из вариантов случайного леса. Для каждого элемента из выборки определяется его мера нормальности или аномальности как среднее арифметическое глубины листьев, которые соответствуют ему.

Смоделируем алгоритм изолированного леса с помощью функции из библиотеки Scikit-learn 3 с параметром максимального количества аномалий 0.15. Так как алгоритм изолированного леса лучше работает на подвыборках из полной выборки 2, такое начальное выделение аномальных ситуаций на этапе предварительной обработки, только улучшает результативность алгоритма из-за понижения плотности аномальных ситуаций в определенных местах.

Так как алгоритм isolation forest самообучающийся алгоритм машинного обучения, то будет произведено измерение точности поиска аномальных пользователей как количество правильно определенных аномальных событий, деленные на общее количество аномальных событий.

Тестирование

На данных размером 131824×871 обучение модели произошло за 2 минуты 12 секунд (Вычисление производилось на Intel i7-4700HQ). Для сравнения вместо алгоритма Isolation Forest была использована альтернативная модель машинного обучения без учителя “локальный уровень выброса”⁴, результаты тестирования можно увидеть в Таблице 2.

Таблица 2. Итоги по моделям Isolation Forest и Local Outlier Factor

	Isolation Forest	LOF (Local Outlier Factor)
Точность определения аномальных событий	0,89	0,336
Время работы алгоритма (секунд)	132	11250,48

Заключение

В работе был предложен алгоритм поиска аномальных событий в журналах событий nginx на основе выявления аномальных событий на этапе предварительной обработки и алгоритма Isolation Forest. Данный подход показал достаточно хороший результат, как по времени работы, так и по точности определения аномальных событий. Так как алгоритм, изолированный леса имеет линейную сложность 2 получается, что алгоритм очень хорошо подходит для использования его при обработке больших данных. Предложенный подход с реализацией метода Isolation Forest не требует больших вычислительных мощностей, что позволяет его использовать на производственном оборудовании.

Алгоритм Isolation Forest показал лучше результат, чем LOF как в точности определения, так и по времени работы, что подтверждает его преимущество в использовании.

Список литературы

1. He P. et al. An evaluation study on log parsing and its use in log mining //2016 46th annual IEEE/IFIP international conference on dependable systems and networks (DSN). – IEEE, 2016. – С. 654-661.
2. Liu F. T., Ting K. M., Zhou Z. H. Isolation Forest //2008 eighth IEEE international conference on data mining. – IEEE, 2008. – С. 413-422.
3. Scikit-learn // Официальный сайт Scikit-learn URL: <https://scikit-learn.org/stable/index.html> (дата обращения: 24.04.2022).
4. Breunig M. M. et al. LOF: identifying density-based local outliers //Proceedings of the 2000 ACM SIGMOD international conference on Management of data. – 2000. – С. 93-104.
5. Geoip2 // Maxmind URL: <https://dev.maxmind.com/geoip> (дата обращения: 13.02.2021).

ОБНАРУЖЕНИЕ ПОДДЕЛЬНЫХ УЧЕТНЫХ ЗАПИСЕЙ В СОЦИАЛЬНОЙ СЕТИ ВКОНТАКТЕ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Сюракишина А.В.

*Национальный исследовательский университет
«Высшая школа экономики»,
департамент электронной инженерии
МИЭМ НИУ ВШЭ*

Аннотация

Данное исследование направлено на выявление поддельных аккаунтов и ботов в социальной сети «ВКонтакте» методами машинного обучения. В результате работы создан уникальный набор пользовательских данных и построена модель классификации, способная выявлять поддельных профилей с высокой точностью. Полученное решение поможет повысить безопасность пользователей в социальной сети ВКонтакте и снизить риск ее использования.

Введение

В настоящее время более половины населения мира является активным пользователям социальных сетей. Люди используют данные сервисы, чтобы коммуницировать друг с другом, делиться новостями и даже вести свой собственный бизнес. Постоянно растущая популярность социальных сетей привела к серьезной проблеме – отсутствию гарантии, что человек, владеющий аккаунтом, реален. Вручную проверить профили всех зарегистрированных пользователей в крупных социальных сетях практически невозможно.

Одним из решений данной проблемы является разработка модели машинного обучения [1], способной классифицировать пользователей на две группы: настоящие и поддельные.

Цель исследования – увеличение безопасности использования пользователями социальной сети ВКонтакте. В первом разделе исследования представлены признаки, наиболее характерные для поддельных аккаунтов. Во втором описывается процесс сбора и обработки данных пользователей социальной сети ВКонтакте. В третьем разделе проведен первичный разведочный анализ данных [2]. Четвертая глава посвящена созданию моделей машинного обучения для целей исследования и анализу результатов их работы.

Признаки поддельных аккаунтов

Фейковая страница в социальных сетях – это аккаунт, персональная информация которой не соответствует реальным данным ее владельца. В некоторых случаях такие страницы созданы реальными людьми, цель которых – скрыть свою персональную информацию. В других – компьютерными программами [3], способными имитировать поведение реальных пользователей. Второй тип аккаунтов можно считать наиболее опасным, так как целью их создания является хищение персональных данных других

пользователей, а также создание искусственной рекламы или ложных новостей.

В результате проведенного анализа, экспертным методом удалось выявить наиболее распространенные признаки фейковых страниц. Стоит отметить, что представленные далее факторы относятся исключительно к первичному визуальному анализу страниц пользователей.

В большинстве случаев дата регистрации поддельных аккаунтов относительно недавняя, так как они часто блокируются из-за жалоб пользователей. Реальный пользователь с большей вероятностью активно пользуется социальной сетью в течение длительного времени. Выявить фактор того, создан ли аккаунт недавно, можно путем анализа активности пользователя на своей странице, а именно: даты загрузки фотографий, даты размещения записей и т.д.

Еще одним весомым признаком фейковой страницы является отсутствие активности друзей на странице пользователя, причем местоположение, указанное для друзей пользователя, имеет сильный разброс по городам.

Также выявлено, что на страницах поддельных аккаунтов чаще отсутствует информация о жизни пользователя: записей о его новостях или мыслях, образовании, статусе, месте работы и т.п.

Выявленные признаки не позволяют со стопроцентной уверенностью распознать фейковый аккаунт, но помогают собрать необходимые для исследования данные.

Сбор и обработка данных

В настоящем исследовании, на данный момент, собрано 126 аккаунтов, 55 из которых поддельные, причем 49 из них являются ботами, а оставшиеся 6 принадлежат реальным пользователям, намеренно скрывшим свои персональные данные. Информация с данных страниц получена с помощью интерфейса VK API [4].

В результате первичной обработки данных удалены все приватные и заблокированные профили. Все данные, представленные в текстовом формате приведены к числовому путем их замены на длину текста или на признак наличия такой информации у пользователя.

Идентификатор пользователя социальной сети Вконтакте изначально представляется в виде порядкового номера зарегистрировавшегося, но, по желанию, его можно изменить, поэтому дополнительно создан признак количества числовых символов в идентификаторе профиля к его длине [5].

Фактор с именем в профиле заменен на числовой токен благодаря частотному анализу. Таким образом, чем имя более распространенное, тем меньше значение этого токена.

Целевая переменная представлена в бинарном виде и равна нулю, если профиль реального человека, и единице в противном случае.

Перечень и описание полученных признаков представлены в таблице 1.

Таблица 1. Описание признаков

Признак	Описание	Тип
id	Идентификатор пользователя	integer
has_photo	Присутствует ли фотография профиля	boolean
has_mobile	Известен ли номер мобильного телефона пользователя	boolean
about_length	Длина содержимого поля «О себе» из профиля	integer
has_movies_info	Содержится ли в поле «Любимые фильмы» информация	boolean
has_music_info	Содержится ли в поле «Любимая музыка» информация	boolean
can_write_private_message	Можно ли пользователю отправить личное сообщение	boolean
has_external_site	Указан ли в профиле внешний сайт пользователя	boolean
followers_count	Количество подписчиков	integer
has_occupation	Указан ли текущий род занятий пользователя	boolean
has_relation	Указано ли семейное положение	boolean
has_personal	Есть ли информация о полях из раздела «Жизненная позиция»	boolean
relatives	Указаны ли родственники	boolean
is_no_index	Индексируется ли профиль поисковыми сайтами	boolean
has_skype	Указаны ли данные о профиле в skype	boolean
has_twitter	Указаны ли данные о профиле в twitter	boolean
has_partner	Указан ли другой пользователь в поле «Семейное положение»	boolean
username_ratio	Доля числовых значений в id пользователя	float
first_name_token	Числовой токен имени пользователя, показывающий частоту встречаемости данного имени в конкретной таблице данных	integer
fake	Является ли аккаунт поддельным	boolean

Разведочный анализ данных

Прежде чем приступить к построению модели, целесообразно проверить, насколько выявленные признаки важны для исследуемой задачи классификации. Для этого воспользуемся алгоритмом случайного леса, благодаря которому посчитается важность признаков [6]. Результаты проиллюстрированы на рисунке 1.

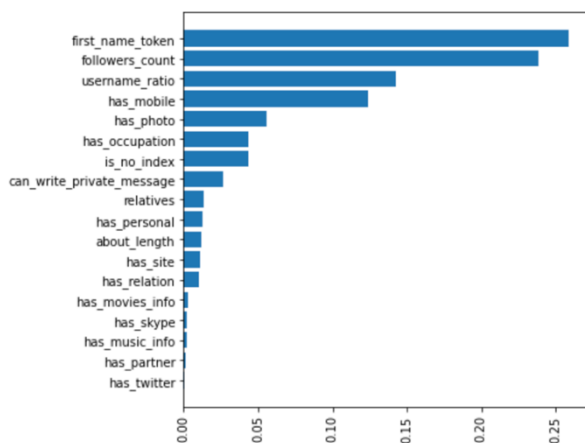


Рис. 1. Отбор признаков с помощью случайного леса

В итоге, не оказались информативными признаки о наличии у пользователя аккаунта в skype или twitter, о факте указания профиля партнера, любимой музыки и любимых фильмов. Это объясняется тем, что доля людей, отмечающих данную информацию у себя в профилях, в наших данных слишком мала. Числовой токен имени пользователя, напротив, показал хорошую корреляцию с целевой переменной. На рисунке 2 можно отчетливо проследить данную зависимость.

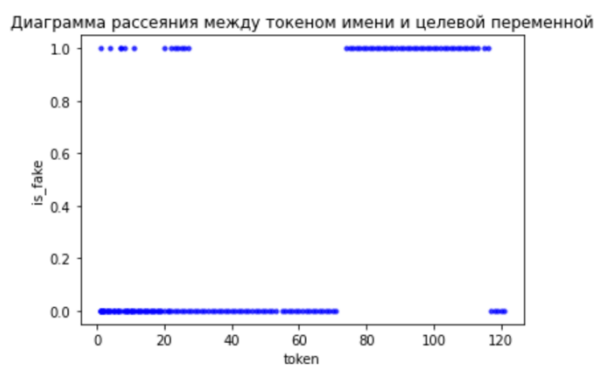


Рис. 2. Диаграмма рассеяния между токеном имени и целевой переменной

Таким образом, чем менее распространено имя, тем выше вероятность того, что аккаунт пользователя окажется поддельным.

Выбор и обучение модели

Завершающий этап работы посвящен изучению работы четырех популярных моделей для классификации: случайного леса, логистической регрессии, градиентного бустинга и метода опорных векторов. По результатам оценки алгоритмов по метрике

ассирусу (1), лучшие результаты показал случайный лес с параметрами: максимальная глубина дерева = 5, число деревьев = 100 (рисунок 3).

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad (1)$$

где TP – доля правильных положительных предсказаний; TN – доля правильных отрицательных предсказаний; FP – доля неправильных положительных предсказаний; FN – доля неправильных отрицательных предсказаний.

```
Model: RandomForestClassifier
train_score: 1.000
validation_score: 0.898
-----
Model: LogisticRegression
train_score: 0.919
validation_score: 0.889
-----
Model: GradientBoostingClassifier
train_score: 1.000
validation_score: 0.886
-----
Model: SVC
train_score: 0.967
validation_score: 0.843
-----
```

Рис. 3. Сравнение моделей

Выбранная модель допустила только одну ошибку на тестовой выборке, приняв поддельный аккаунт за настоящий (рисунок 4). Таким образом, итоговая ассируса модели = 0,97.

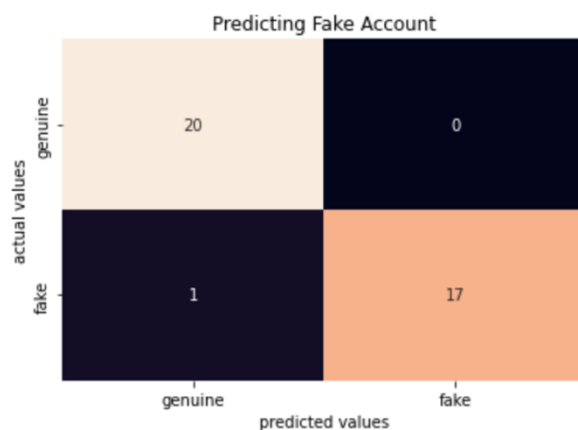


Рис. 4. Матрица ошибок случайного леса

Заключение

В описанном исследовании для борьбы с постоянно растущим числом поддельных аккаунтов предложена модель, которую можно использовать для автоматической детекции фальшивых аккаунтов, с целью их удаления до того, как они причинят кому-либо вред. На данном этапе, на относительно малой выборке модель случайного леса уже показала данную возможность для социальной сети «ВКонтакте».

В дальнейшем, планируется увеличить набор данных, добавить признаки, которые были выявлены в ходе экспертного анализа поддельных профилей и провести эксперименты с помощью методов глубокого обучения.

Список литературы

1. Сюрakшина А.В., Перов А.А. Обнаружение поддельных учетных записей в социальных сетях с использованием алгоритмов машинного обучения // тезисы I Новосибирского Международного форума «Большие данные», 2022 [В печати].
2. П. Брюс, Э. Брюс. 1. Разведочный анализ данных // Практическая статистика для специалистов Data Science. — СПб.: БХВ-Петербург, 2018. — С. 19—58. — 304 с.
3. Кузьмин В.Н., Менисов А.Б., Шастун И.А. Методический подход к выявлению ботов в социальных сетях на основе специального объединения классификаторов // Бизнес-информатика. — 2020. — №3. — С. 54-66.
4. Методы API [Электронный ресурс] // VK для разработчиков. URL: <https://dev.vk.com/method/users.get> (дата обращения: 15.01.2023);
5. Instagram fake spammer genuine accounts [Электронный ресурс] // Kaggle. URL: <https://www.kaggle.com/datasets/free4ever1/instagram-fake-spammer-genuine-accounts> (дата обращения: 23.01.2023).
6. Random Forest for Feature Importance [Электронный ресурс] // Medium. URL: <https://towardsdatascience.com/random-forest-for-feature-importance-ea90852b8fc5> (дата обращения: 23.01.2023).

АНАЛИЗ И ОПРЕДЕЛЕНИЕ РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ С ПОМОЩЬЮ ПРОГРАММНОГО МОДУЛЯ ОЦЕНКИ

Маслова М.А.

Севастопольский

государственный университет,

Ростовский государственный

экономический университет (РИНХ)

Аннотация

В работе проводится анализ существующих методов и стандартов для оценки рисков информационной безопасности, разрабатывается программный модуль для анализа и оценки рисков ИБ, содержащий большую базу данных входных параметров для более детального и точного анализа с последующим вычислением на сходимость, и выдачей конкретных результатов по анализу, уменьшению, устранению, своевременному предупреждению возможных рисков ситуаций.

Введение

С появлением данных их различные риски по краже, копированию, модификации становятся все более актуальнее. Появляются не только уязвимости, относящиеся к человеческому фактору, продаже данных или их краже, но и развитие цифровых технологий приводит к постоянным рискам и уязвимостям, совершаемым 24/7 с использованием техники, интернет, сетей и т.д.

Оценка рисков проводится в основном на качественном или количественном уровне в зависимости от требований заказчика, экономической целесообразности и возможностей ее проведения [1].

Необходимо проанализировать существующие методы для оценки рисков информационной безопасности, выбрать часто используемые, выделить положительные и отрицательные стороны методов, сформировать свою базу входных и выходных параметров для дальнейшего внедрения и реализации программного модуля, для анализа и оценки рисков информационной безопасности применяемых в организациях.

Анализ методов оценки рисков информационной безопасности

Анализ и оценка рисков информационной безопасности проводится следующими общеизвестными методами: качественным, количественным, комплексным, которые были описаны и проанализированы [6]. Они выполняются в основном в организациях руководителями, рабочими, компетентными в данной области, методом мозгового штурма или привлекаются экспертами с дальнейшим получением различных форм оценки: в виде четких действий и рекомендаций к ним, форм отчетов (различного вида в зависимости от необходимых требований или используемого метода, или стандарта).

Были рассмотрены классические методы анализа рисков и проанализированы наиболее часто используемые и основополагающие методы, стандарты и методики анализа и оценки рисков информационной безопасности, а также разделены на группы по методам их оценки: качественного, количественного, комплексного (см. табл. 1) [2 - 5].

Таблица 1. Методы оценки рисков информационной безопасности

Качественный	Количественный	Комплексный
– FRAP	– BS 7799	– ISO/IEC
– Risk IT	– RiskWatch	27001
– CORAS	– ГОСТ Р	– OCTAVE
– MSAT	ИСО/МЭК ТО 7	– СТО БР
– MOF risk management	– CRAMM	ИББС
		– ГРИФ
		– Microsoft

По каждому методу были выявлены используемые входные параметры и разделены на положительные и отрицательные их характеристики [6]. Также самое было сделано с выходными параметрами рассматриваемых методик. Также рассмотрены ин-

теллектуальные методы анализа рисков и саморазвивающихся систем защиты [7].

Составить одну стандартную схему построения модели для анализа и оценки рисков не представляется возможным, так как в любой организации есть свои требования, ценности и необходимые параметры [8, 9]. Но общую схему можно составить. Для этого сначала необходимо определить потребности и выделить основные главные параметры для оценки:

- выявить главные активы для предприятия: человек-ресурсы, время, финансовая составляющая, управление, защита ПК, защита сети и т.д.;
- использование и применение в организации нормативных актов, законодательной базы и др. регламентирующих документов;
- выявление определения метода оценки – качественного, количественного, комбинированного;
- детальность и важность проработки возможных угроз и уязвимостей;
- необходимость выявления базовых направлений для управления и уменьшения рисков информационной безопасности;
- получение результата в виде форм, отчетов, списков или деталей;
- быстрота, компетентность и точность проведения экспертизы.

Далее, в зависимости от выявленных параметров, необходимо выбрать подходящий под требования определения метод, который будет удовлетворять выше выявленным параметрам и не затруднять работу, понятен, получен в удобной для этого форме, финансово оправдан. Он должен помогать уменьшать и предупреждать риски, т.е. необходимо учитывать и прорабатывать большое количество параметров, что занимает множество времени, затрат и не дает полного понимания.

В разрабатываемом программном модуле представлена большая база данных входных параметров, которые позволяют в большей степени оценить и проработать детально различные виды рисков и факторов, влияющих на них. С помощью полученных данных были выявлены основные отрицательные характеристики методов и разработаны действия по их устранению и улучшению, которые должны соответствовать следующим требованиям:

- возможность добавлять, улучшать и модифицировать базу входных данных;
- простота в управлении;
- иметь русскоязычный и понятный интерфейс;
- возможность работы с программой в любом часовом поясе в режиме 24/7;
- быть финансово выгодной;
- содержать различные дополнительные важные параметры об организации, для которой будет проводиться экспертиза;
- иметь большую базу экспертов различной квалификации и направленности (постоянно увеличивающуюся);
- выдавать четкие рекомендации, как по анализу, предупреждению, так и по устранению рисков;

- возможность хранения отчетов в течение некоторого времени в личном кабинете;
- выдавать четкую и понятную информацию;
- иметь функции управления, восстановления;
- иметь постоянные обновления.

Т.е. необходимо, чтобы разработанный программный модуль был как можно лаконичнее, проще, учитывая различные влияющие на него параметры с постоянной модернизацией.

Разработка программного модуля анализа и оценки рисков информационной безопасности

Представлены параметры по каждому методу с используемыми в нем данными, полями и размерностями. После определения для каждого из них данные анализируются и составляется общая таблица с одинаковыми, а также разными полями данных. Далее это все формируется в общую базу и по полученным данным эксперты будут проводить экспертизу, определяя значимость для определенной организации, выставляя оценки каждому параметру от 0 до 10, как не важный и очень важный соответственно.

На следующем шаге выполняется расчет сходимости данных с помощью метода Дельфи: определения значения по каждому элементу, которые зависят от важности элемента, компетентности и опыта эксперта по данному вопросу. Методика расчета представлена в работе [10].

Для того, чтобы эксперты могли как можно лучше понять потребности организации, при регистрации клиенты указывают информацию о себе с дополнительным описанием (см. рис. 1)

Рис. 1. Окно регистрации клиентов для подачи заявки на экспертизу

После просчета сходимости параметров и ранжирования их по критерию важности, программа выдает файл с рекомендациями, которые разбиты на важные, средние и не важные в формате Excel. Данный файл отсылается на почту клиенту и хранится в базе. В случае утери файла, клиент может зайти в свой личный кабинет и скачать его там еще раз.

Заключение

Разрабатываемый программный модуль будет содержать большую дополняемую базу данных, с помощью которой будет проводиться экспертиза анализа и оценки рисков ИБ экспертами, автоматический расчет сходимости оценки, выдача рекомендаций от

самого значимого до менее значимого с градацией по важности, разделенные на три категории (значимые, средняя значимость, не значимые параметры/активы).

Благодарность. Работа выполнена в рамках Соглашения от 30.06.2022 г. № 40469-21/2022-к.

Список литературы

1. Залесская, Л. Н. Управление рисками проектов, концепции рисков, количественный и качественный анализ рисков / Л. Н. Залесская, В. В. Митина, Е. А. Евдокимова // Актуальные исследования. – 2020. – № 9-2(12). – С. 78-81.
2. Обзор методик и анализ рисков информационной безопасности, информационных систем предприятия [Электронный ресурс]: база данных. – Режим доступа: <https://cyberleninka.ru/article/v/obzor-metodik-analiza-riskov-informatsionnoy-bezopasnosti-informatsionnoy-sistemy-predpriyatiya>.
3. Microsoft Security Assessment Tool 4.0 [Электронный ресурс]: база данных. – Режим доступа: anti-malware.ru
4. Пугин, В. В. Обзор методик анализа рисков информационной безопасности информационной системы предприятия / В. В. Пугин, О. Ю. Губарева // Т-Comm: Телекоммуникации и транспорт. – 2012. – Т. 6. – № 6. – С. 54-57.
5. Абдуллаев, А. Р. Риски в инновационных проектах: причины появления, интегральные риски, экспертиза проектов с учетом рисков / А. Р. Абдуллаев, Л. А. Мыльников, Е. Е. Васильева // Экономический анализ: теория и практика. – 2012. – № 40(295). – С. 41-49.
6. Маслова, М. А. Сравнительный анализ методов оценки рисков информационной безопасности, основанных на стандартных и интеллектуальных подходах / М. А. Маслова, Е. Н. Тищенко // Проблемы проектирования, применения и безопасности информационных систем в условиях цифровой экономики: материалы XIX Международной научно-практической конференции, Ростов-на-Дону, 28–29 октября 2019 года / Ростовский государственный экономический университет (РИНХ). – Ростов-на-Дону: Ростовский государственный экономический университет "РИНХ", 2019. – С. 211-215.
7. Мунтян, М. М. Формализация процессов обработки информации с использованием методов интеллектуального анализа в системах мониторинга рисков информационной безопасности / М. М. Мунтян // Инженерные кадры - будущее инновационной экономики России. – 2022. – № 1. – С. 453-459.
8. Ibrahimova, A. N. The factor of the risk and risk management in information security / A. N. Ibrahimova // Journal of Actual Problems of Jurisprudence. – 2021. – Vol. 100. – No 4. – P. 97-106. – DOI 10.26577/JAPJ.2021.v100.i4.12.
9. Биктимерова, Э. Ю. Риски предприятия как составная часть рисков / Э. Ю. Биктимерова // NovaInfo.Ru. – 2016. – Т. 1. – № 56. – С. 201-206.
10. Маслова, М. А. Анализ, применение и модификация метода Дельфи / М. А. Маслова // Вестник УрФО. Безопасность в информационной сфере. – 2022. – № 2(44). – С. 25-29. – DOI 10.14529/secur220203.

ИНТЕГРАЦИЯ РАСПРЕДЕЛЁННЫХ РЕЕСТРОВ В ВЫСОКОНАГРУЖЕННЫЕ ТУРИСТИЧЕСКИЕ СИСТЕМЫ

*Головинский С.А., Симоненков Н.В.,
Воробьев М.О., Пилипенко А.С.
Севастопольский государственный университет*

Аннотация

В работе обоснована целесообразность применения системы распределённых реестров в сфере туризма. Проанализирована интеграция распределённых реестров в высоконагруженные системы. Разработана концепция системы поиска оптимальных туристических маршрутов на основе системы блокчейн.

Введение

В настоящее время существующие туристические сервисы имеют различные проблемы с отладкой нагрузки на механизмы, испытывая трудности с содержанием серверной и аппаратной части. Это сказывается и на общей безопасности, так как неустойчивая система влияет на качество хранения информации, а нестабильность может привести к утечке или иным последствиям. Появление сервисов, использующих мало защищённые системы, может негативно сказываться на безопасности проведения сделок и достоверности в рамках выбора туристического маршрута. Текущий туристический рынок РФ не предлагает платформы с использованием системы распределённых реестров, которая бы повышала качество оказываемых услуг и обеспечивала достоверность операций.

Разрабатываемый алгоритм основан на том, что предоставляет пользователям и операторам услуг платформу, способную к интегрированию в различные высоконагруженные системы с возможностью выбора различных туристических услуг с выбором маршрутов и их вариаций с системой отбора критериев пользователя. Модель нацелена на обеспечение максимального удобства пользователей в рамках выбора комплексного обеспечения услуг и гарантий достоверности операций с использованием технологий блокчейн.

В данной работе рассматривается концепция распределённого приложения для поиска оптимального туристического маршрута с интеграцией в различные системы.

Актуализация распределённых реестров в туристической сфере

Блокчейн — это технология обработки данных, в основе которой можно выделить следующие основные принципы: структура хранения данных представляет собой выстроенную по определенным правилам цепочку блоков, содержащих информацию; каждый блок цепочки связан с соседними блоками криптографическими методами; сами по себе блоки и содержащаяся в них информация являются общедоступными; копии цепочки блоков хранятся на разных компьютерах [1].

Работа распределённой системы основана на ряде принципов, обеспечивающих её надёжное функционирование [2]:

- хранение информации о транзакции между пользователями и проверка её достоверности;
- запись транзакций или блоков в распределённую базу данных;
- наличие у оператора системы копии базы данных, соединённой с носителями сети;
- верификация записи при проведении транзакции и добавление её к блокчейну.

Однако, существующие методы защиты от фальсификации как часть систем информационной безопасности нуждаются в развитии, так как создаётся новая аппаратура и новые протоколы безопасности. Уязвимость различных систем хранения данных подтверждается недавним инцидентом, возникшим у одного из главных онлайн сервисов по поиску и покупке туров Level.Travel. В данной компании не используется блокчейн-система, предпочитается «классический» облачный вариант. Одной из преступных хакерских группировок NLB был осуществлен взлом данных компании. Попавшая в международную сеть информация содержит огромное количество пользовательских данных, которые включают в себя более 720 тысяч уникальных номеров телефонов пользователей и более миллиона адресов электронных почт, подкреплённые ФИО пользователей, датой рождения, серий и номером паспортов, идентификаторов различных социальных сетей, IP-адресов [3]. Поэтому, важно продолжать развитие методов защиты от фальсификации, включая использование технологии блокчейн в различных сферах, включая индустрию туризма и путешествий.

Особенности использования технологии блокчейн заключаются в обеспечении гарантий пользователям на своевременное выполнение сервисных услуг и хранение информации о подтверждении транзакций в блоках. Оптимальность дальнейшего разрабатываемого алгоритма оценивается по четырём критериям: доступность, безопасность, ориентированность на клиента и эффективность. Технология блокчейн при использовании в туристической сфере обладает рядом преимуществ:

- повышенная безопасность;
- прозрачность операций;
- смарт-контракты;
- улучшенные программы лояльности;

В целом, использование технологии блокчейн в туризме имеет потенциал для повышения безопасности, эффективности, прозрачности и удовлетворенности клиентов в этой отрасли.

выкуп

Алгоритмизация продукта

Проведённые исследования позволили разработать схему работы продукта. Схема представлена на рис. 1.



Рис. 1. Схема алгоритма работы распределённой системы по поиску туристических маршрутов

Схема структурированно описывает выполнение основных функций продукта. На практике, пользователю предлагаются различные критерии отбора, на основе которых механизм машинного обучения составляет маршрут, внося его в структуру блокчейн. В ходе прохождения данного маршрута модель прогнозирования, рассчитывая действия пользователя, а также оценок его шагов, в режиме реального времени предлагает ему альтернативный или дополнительный маршруты к уже созданному.

Интегрированная в высоконагруженную систему технология блокчейн выполняет функцию платформы, способной располагать в себе различные туристические услуги, предоставляемые операторами. В блокчейн предлагается хранить пользовательскую информацию о получаемых туристических услугах, персональные данные, а также маршруты пользователя. Вопрос хранения персональных данных является ключевым, так как для его реализации применяются наиболее современные алгоритмы симметричного и асимметричного шифрования. При необходимости обработки биометрических данных может найти применение метод маскирующего сжатия на основе модели взвешенной структуры изображения [4]. В целом, система распределённых реестров применяется для хранения пользовательской информации о получаемой услуге (функция подтверждения достоверности), алгоритм выбора критериев пользователя, основанный на использовании машинного обучения с моделью прогнозирования на основе TensorFlow. Алгоритмы транзакций — это набор правил, которые позволяют обеспечить целостность информации в цепочке блоков.

На основе описанного алгоритма можно разработать программный продукт, обеспечивающий взаимодействие клиента и виртуального туроператора, предоставляющего различные услуги. Система гарантирует невозможность ошибочно или ложно из-

менить данные.

Главной задачей выступает интеграция технологии блокчейн в высоконагруженные системы. В отличие от централизованных систем, блокчейн – распределённый реестр, работающий без участия человека.

Внедрение блокчейна в высоконагруженные системы позволяет избавить их от ненужных вычислений и операций, используя децентрализацию. Блокчейн берёт на себя определённую часть нагрузки, оптимизируя работу и снижая затраты.

Заключение

Осуществив интеграцию распределённых реестров в высоконагруженные туристические системы, её владелец получает возможность оптимизировать нагрузку на свою систему за счёт особенностей архитектуры блокчейн. Важно не «перегружать» блокчейн, а лишь отдавать ему необходимую роль в системе, так как его эффективность существенно зависит от его места в архитектуре [5].

Таким образом, в результате проведенных исследований разработан алгоритм, позволяющий, реализовать платформу, способную на основе системы блокчейн являться гарантом достоверности сохранности пользовательских данных и информации о проведенной операции. Уже сейчас можно говорить о том, что блокчейн может существенно изменить туристическую сферу, сделав ее более прозрачной

для всех участников этого рынка.

Список литературы

1. Козин И.С. Метод обеспечения безопасной обработки персональных данных на основе применения технологии блокчейн // Научно-технический вестник информационных технологий, механики и оптики. 2019. №5. URL: <https://cyberleninka.ru/article/n/metod-obespecheniya-bezopasnoy-obrabotki-personalnyh-dannyh-na-osnove-primeneniya-tehnologii-blokcheyn> (дата обращения: 15.04.2023).
2. Никитина А.А Тищенко С.В. Технологии блокчейн-инновационный прорыв в туризме // Проблемы экономики и юридической практики. - 2018. - №2. - С. 218-220.
3. Утечка пользовательских данных Level.Travel [Электронный ресурс]. — Режим доступа: <https://xakep.ru/2022/12/15/level-travel-leak/> (дата обращения 04.01.2023).
4. Беззатеев С.В., Волошина Н.В. Маскирующее сжатие на основе модели взвешенной структуры изображения // Информационно-управляющие системы. 2017. № 6. С. 88–95.
5. Высокоуровневый блокчейн в mission critical системе // HighLoad SPB URL: <https://highload.ru/spb/2022/abstracts/8726> (дата обращения: 16.01.2023).

СОДЕРЖАНИЕ

<u>Секция 1. «Математика и компьютерное моделирование»</u>	5-61
Буров Н.А. Двухпалубная структура пограничного слоя в трехмерной задаче обтекания малой неровности на поверхности пластины	5-7
Юдина А.В. Минимальное покрывающее множество как инструмент оптимального коллективного выбора	7-11
Мозоленко В.К. Численное исследование критического поведения в модели Блюме-Капеля методом прямой оценки энтропии	11-13
Загороднев Д.Г. Кобякова А.А. Пивоваров А.С. Попов Л.Д. NLP-методы детектирования существенных изменений в программном коде математических моделей	13-17
Румянцева С.В. Осциллирующее расщепление нижних энергетических уровней квадратичного оператора на алгебре $SU(1,1)$	17-19
Семион А.А. Дифференциальная игра перехвата и защиты цели	19-21
Федченко А.С. Свойства агрегированной квазигазодинамической системы уравнений многокомпонентной газовой смеси	21-24
Константинова Е.П. О плотности распределения углов в сечениях многогранника случайными плоскостями	24-27
Масликов Д.Н. Шарова А.И. О свойствах одной стохастической модели динамики мнений с ограниченным радиусом доверия	27-30
Суховерхова Д.Д. Нейросетевой анализ фазовых переходов в классических моделях Изинга	30-33
Антонов Д.А. Моделирование пространственно-распределенных эволюционных игр со случайной сменой стратегий	33-35
Денисова А.И. Методика расчета парциальных давлений системы: этаноламинового раствора – H_2S-CO_2	35-39
Гафурова Д.Р. Исследование методов численного интегрирования уравнений круговой ограниченной задачи трёх тел	39-41
Миколаенко В.В. Построение модели деформационного поведения сверхпластических материалов	42-45
Зеленецкая Д.Н. Пучков Д.В. Девянин П.Н. Об автоматической генерации Си-кода на основе Event-B спецификаций	45-48

Михайлова С.О. Петухова Е.С. Асимптотика решений уравнений на сетках в дискретном времени	48-50
Гаража И.А. Барина С.А. Регуляризованный фильтр Калмана в приложении к Чандлеровскому движению полюса	51-53
Рахель М.А. Построение и обоснование асимптотики фундаментального решения для вырождающегося одномерного параболического уравнения с малым параметром	54-56
Пиле Я.Э. Связанные состояния в модели Хаббарда на лестничных решетках	56-59
Гуськова М.С. Использование модели со скоростью звука $\sqrt{2/3}$ в методе решеточного уравнения Больцмана на примере движения частицы в потоке Пуазейля	59-61
<u>Секция 2. «Информационно-коммуникационные технологии»</u>	62-172
Мотявин М.В. Чихватова А.А. Поиск изображений на мобильном телефоне (ОС Android) с использованием методов машинного обучения	62-64
Константинов А.С. Сидоров И.В. Разработка программы генерации изображений с использованием комбинации CNN, GAN и диффузионных методов машинного обучения	65-67
Селин М.П. Исследование оптимальных графов циркулянтной топологии	68-70
Овчинникова А. А. Автоматизация верификации СБИС с помощью методов поиска ошибки	70-71
Рзаев Э.Р. Рыжов А.Р. Исследование иерархических топологий для проектирования сетей на кристалле	72-75
Анджушева М.М. Бичурина Е.П. Методика создания датасета автомобильных аварий	75-77
Баканов Г.П. Подкопаева П.О. Цыгикало П.Н. Савинов А.И. Система автоматизированного расчета стационарных тепловых режимов интегральных микросхем	78-81
Киреенко Л.В. Разработка методики и алгоритмов автоматизированного тестирования графических приложений на аппаратно-программной платформе «Эльбрус»	82-84
Волков П.О. Обзор аналитических методов оценки надежности технических средств	84-86
Смоленчук И.К. Нижегородов М.А. Разработка алгоритма для создания датасета на русском языке для задачи преобразования текста в речь	87-89
Гаранина А.Е. Иванова Е.А. Формализация методики оценивания состояния программных проектов по критериям прогресса и здоровья проекта на базе стандарта OMG Essence	89-92

Горбачева В.Ю. Жукова М.А. Зарубина Т.В. Ильинская Е.В. Палуха Д.В. Разработка системы предотвращения аварийных ситуаций при 3D печати	92-94
Хачатрян А.А. Емельянова А.К. Михайловская А.Г. Нгуен А. Разработка параметризованного 3D принтера	94-96
Цуркан А.А. Старилова Е.В. Нигматулин Н.Р. Сватухин А.М. Жибуртович Е.М. 3D-реконструкции и моделирование исторического контента Дома Дурасовых	97-100
Бородин Н.Ю. Пономарев А.С. Таржанов Т.В. Разработка САПР для высокоуровневого моделирования сетей на кристалле	100-102
Степанянц В.Г. Андрончев А.Д. Шабалин Е.А. Добжанский Е.К. Проектирование среды моделирования каналов и сетей связи подключенного транспорта	102-105
Маршутина Е.Н. Разработка модулей системы управления и узлов мультифункционального стенда для исследования кристаллических материалов	106-108
Волгина О.А. Нечаева П.А. Моделирование гетерогенных информационных взаимодействий пользователей социальной сети	108-110
Козырева Д.Д. Алгоритмы искусственной сегментации и генерации наборов данных для сферы ручного производства	111-113
Граченко Е.Н. Демодуляция многопозиционных сигналов с использованием нейронной сети	114-116
Валиахметова К.А. Методика прогноза изменения характеристик базовой станции в результате модернизации оборудования	116-118
Казакова В.Е. Разработка устройства для самоконтроля зрения человека	119-121
Козлов А.С. Шаповалов М.В. Энергоэффективный метод передачи данных в гетерогенной сети интернета вещей по спутниковым каналам связи	122-124
Ан Д.А. Щербаченко Н.А. Будаев М.В. Карманова Н.Д. Мосолов А.М. Разработка веб-интерфейса для стенда удаленной работы с оборудованием	125-126
Ранасингхе М.П.К.К. Кантута З.Д. Подготовка образа, тестирование и разработка утилиты для одноплатного компьютера Raspberry PI 4 на базе ос ALT Linux	126-129
Угодников А.А. Антипов М.М. Программный эмулятор устройств Интернета вещей "Virtual IoT Lab"	129-131
Образцов О.А. Алгоритм множественного доступа с контролем несущей и избеганием коллизий (CSMA/CA) сети Meshtastic	131-134

Евтушенко Л.Г. Экспериментальное исследование наличия установочных трасс заданной длины для неинициальных автоматов	134-136
Минубаев М.И. Библиотека для синхронизации реплицированных микросервисов	136-139
Бандюков А.А. Метляков А.А. Мягкое декодирование БЧХ кодов	139-140
Котов А.В. Кадыргулов Т.А. Создание образовательной среды для изучения искусственного интеллекта в рамках 2D платформера	141-144
Конаныхин К.А. Полетаев Д.А. Программный пакет для историков	144-145
Ильин А.Д. Разработка и исследование методов повышения энергетической эффективности граничных устройств в сети IoT	145-147
Моисеев М.В. Галкина М.А. Мирталибов Д.И. Пивоваров А.С. Орлов К.О. Исследование и описание реализации системы автоматической съемки мероприятий разговорного формата	147-150
Боярский А.О. RTSP-Proxu: экспериментальное исследование существующих решений	150-154
Андреев А.Г. Среда моделирования атакующих сценариев алгоритма консенсуса Tangle	154-156
Заякина А.М. Гронин Д.С. Разработка способа создания 3D фона для студии с хромакеем	156-158
Кофман М.О. Лютиков А.С. Обзор подходов к разработке мобильных приложений	159-162
Соколовская К.Е. Курс "Сетевые видеотехнологии"	162-165
Ефремов А.М. Елисеенко А.М. Модульная система управления сетевыми видеокамерами и программными компонентами видеостудий с помощью физических устройств ввода	165-170
Задорожникова А.А. Зотова А.В. Царегородцева Е.И. Система удаленных медицинских осмотров	170-172
<u>Секция 3. «Электроника»</u>	173-282
Пригородов Р.А. Обеспечение устойчивости к механическим воздействиям бортового электронного устройства	173-175
Опекунова А.А. Демидов И.Д. Разработка структурной схемы бесконтактного измерителя вибрации	175-178
Давыдов С.С. Коробейникова Т.И. Разработка макромоделей операционных усилителей с учетом воздействия на них температуры	178-181

Тюрина А.Н. Демочкидов А.Р. Система распознавания борщевика Сосновского на базе беспилотного летательного аппарата	182-184
Зубкова А.И. Разработка цифровых двойников для мощных МОП транзисторов	184-185
Уркунов А.К. Цветков В.Э. Ландер Л.Б. Сафина М.Ф. Разработка структуры программно-аппаратного комплекса для проведения ускоренных испытаний	186-188
Демочкидов А.Р. Тюрина А.Н. Разработка методики уточненной оценки показателей безотказности современных интегральных микросхем	188-190
Артамонов Д.О. Уткин Б.В. Метод передачи информации в условиях воздействия мощных широкополосных помех	191-193
Кобяков В.В. Создание программы учета при схемотехническом моделировании воздействия на работу КМОП схем внутренних факторов старения	193-194
Камаров Л.Ш. Шумкова К.Ю. Создание схемотехнических моделей элементной базы отечественных ОУ	194-197
Клопотов Г.И. Ивлев К.И. Моделирование электро-теплового режима блока управления электрическим двигателем топливного насоса	197-198
Беликов И.Д. Варианты модернизации модуля спецвычислителя бесплатформенной инерциальной навигационной системы	198-200
Филяев А.А. Заводиленко В.В. Павлов И.Д. Ожегов Р.В. Моделирование InGaAs/InAlAs однофотонных лавинных фотодиодов с целью оценки влияния диаметра активной области на уровень темнового счета	200-202
Башкевич С.В. Лаврухин И.Р. Исследование влияния фрактального метаматериала на электродинамические параметры RFID-антенны	202-205
Сетдинов Р.М. Тюрин И.Г. Разработка аналоговой части спектрометрического детектора, регистрирующего ионизирующие излучения	205-208
Минеев А.В. Мкртчян Г.А. Насыров Д.Д. Озорнина М.М. Разработка концепции USB-накопителя с индикатором технического состояния	208-211
Васьков М.О. Генератор аналогового сплайн-сигнала на основе первого и второго интерполятора Ньютона	211-213
Толстикова С.Ю. Муллахметов И.Р. Программное обеспечение для верификации экспериментальных данных по радиационной электропроводности полимеров	213-214
Иконникова Д.М. Кацнельсон А.И. Громова П.И. Ильянова Д.С. Мониторинг потерь радиочастотной энергии мобильной связи в окружающую среду мегаполиса Москвы	215-217

Матинян А.А. Вяхирева П.И. Мельникова Е.А. Редукция численных методов для построения макромоделей эквивалентных электрических схем космических аппаратов	217-219
Ярмиева К.Р. Виртуальные лабораторные стенды по курсу электроники	219-222
Махлуф М.М. Елезов М.С. Измерение временного разрешения разработанного двухканального время-цифрового преобразователя на основе XILINX ZYNQ SOC 7000	223-225
Рахимов Р.М. Разработка структурной схемы и алгоритмов для управления встроенными средствами летательного аппарата при подготовке к вылету или проведению регламентных работ	225-227
Старилова Е.В. Влияние технической надежности электронных изделий на экономическую надежность организации	227-230
Гронин Д.С. Оценка надежности скоростного вычислительного модуля-синхронизатора оптической линии для управления фазовыми модуляторами оптической системы	230-233
Горбунов К.А. Исмаил-Заде М.Р. Измерение и моделирование ВАХ полевых транзисторов с управляющим рп-переходом	233-234
Каунг Мьят Хту Афонин А.Н. Применение вибрационных микрогенераторов для питания беспроводных сенсорных сетей	235-237
Ротай А.Д. Разработка структуры системы сбора и обработки данных о повреждениях дорожного полотна	237-239
Недошивина М.С. Язовская Д.А. Модификация пользовательского интерфейса измерительного комплекса	239-241
Чуйко Д.О. Кретьова А.А. Сравнение способов разработки автоматизированной системы самообслуживания для инвентаризации для библиотеки на основе RFID-меток и QR-кодов	241-245
Макаренко Д.В. Матвеев Д.С. Концепция системы обнаружения источников радиосигнала на базе БПЛ	245-247
Захваткина В.Ф. Дубовикова С.А. Некрасов Ф.И. Система отслеживания человекопотока в реальном времени	247-249
Кунижев И.Р. Обзор и анализ технологий мониторинга парниковых газов	249-251
Казакова В.Е. Хобов А.П. Разработка аппаратной части стенда для проведения массовых лабораторных работ по электронике	251-253
Сергеенко К.А. Дубильер Я.А. Санникова Н.В. Разработка виртуальных моделей измерительных преобразователей	254-256

Кучина А.П. Исследование и разработка методики расчета мониторинга поля, создаваемого конической рупорной антенной	257-258
Саповский И.М. Рахматуллаев Т.Р. Пинчук К.Е. Создание аморфных и аморфно-нанокристаллических металлических сплавов с заданными служебными свойствами	259-261
Зубкова А.И. Пуговкин М.Д. Смирнов Н.В. Исследование оптических устройств с материалами с фазовым переходом для оптической памяти	262-264
Савочкин В.В. Разработка макромодели операционного усилителя с учетом влияния радиации	264-267
Ларионов С.А. Уланова А.В. Применение методов машинного обучения для построения модели МОП-транзистора	267-270
Корчажинская С.А. Баева Э.М. Время релаксации энергии в эпитаксиальных TiN пленках	271-273
Вовк Н.А. Шибалов М.В. Мумляков А.М. Красильников М.И. Тевяшов А.А. Смирнов К.В. Ануфриев Ю.В. Исследование влияния эффекта поля на электрические характеристики структур с унт на основе диода Шоттки	273-275
Калачикова И.В. Получение нанопроволок из железа, исследование их структуры	275-277
Бронников О.А. Гостищев П.А. Лучников Л.О. Саранин Д.С. Нанокристаллический оксид никеля р-типа для перовскитной оптоэлектроники с применением лучевых технологий	278-280
Бадалян Р.К. Кержеманкина К.В. Сафронов М.А. Тыщевский А.Ю. Эфендиев С.В. Метод определения температуры экструзии пластика	280-282
<u>Секция 4. «Информационная безопасность»</u>	283-307
Бурба А.В. Фокина А.И. Применение платформы ТХМ для анализа текстов различного типа	283-285
Аванесян Н.Л. Зенькова В.В. Попов В.А. Выделение неявных пересекающихся сообществ на графе взаимодействия telegram-каналов	285-289
Шелепова А.Н. Воробьев И.А. Исследование применения методов машинного обучения в задаче выявления мошеннических действий в отношении клиентов банка при подтверждении операции	289-292
Малашина А.Г. Исследование энтропии текстов на естественном языке методом составления словарей n-грамм	292-294
Эшкинин А.И. Доманова В.С. Фонарева О.Е. Мартынов А.С. Применение состоятельных мер зависимости к решению задачи обнаружения инсайдера	294-2973

Федоров С.А. Исследование аномальных ситуаций сетевого трафика по данным журналов web-сервера nginx	297-299
Сюракшина А.В. Обнаружение поддельных учетных записей в социальной сети Вконтакте с использованием алгоритмов машинного обучения	299-302
Маслова М.А. Анализ и определение рисков информационной безопасности с помощью программного модуля оценки	302-304
Головинский С.А. Симоненков Н.В. Воробьёв М.О. Пилипенко А.С. Интеграция распределённых реестров в высоконагруженные туристические системы	304-306
СОДЕРЖАНИЕ	307-314

**Межвузовская научно-техническая конференция студентов,
аспирантов и молодых специалистов имени Е.В. Арменского.
Материалы конференции.**



Подписано в печать 12.10.2023 г. Формат 60х84/8. Бумага офсетная №2.
Печать ризография. Усл.печ.л. 39,25. Уч.-изд.л. 35,33. Тираж 100 экз.
ООО "Издательский дом Медиа Пабlishер".