

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»



Новосибирск  
2024

Всероссийская научная  
конференция молодых ученых

# НАУКА ТЕХНОЛОГИИ ИННОВАЦИИ

Сборник  
научных трудов

Часть 2

г. Новосибирск, 04-08 декабря 2023 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

---

---

# НАУКА ТЕХНОЛОГИИ ИННОВАЦИИ

---

---

**XVII Всероссийская научная конференция молодых ученых  
г. Новосибирск, 04-08 декабря 2023 г.**

**Сборник научных трудов  
в 11-и частях**

**Часть 2**

НОВОСИБИРСК  
2024

УДК 62(063)  
ББК 72.5я431  
Н34

**Н34 НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ:** XVII Всероссийская научная конференция молодых ученых (г. Новосибирск, 04-08 декабря 2023 г.) // Сборник научных трудов в 11 ч. / Под ред. Соколовой Д.О. – Новосибирск: Изд-во НГТУ, 2024.

ISBN 978-5-7782-5128-1  
Часть 2: – 201 с.  
ISBN 978-5-7782-5130-4

В сборнике опубликованы материалы научных исследований, представленных на XVII Всероссийской научной конференции молодых ученых.

В части 1 сборника представлены материалы по направлению «Информатика, автоматика, вычислительная и измерительная техника». В части 2 сборника представлены материалы по направлению «Информационные технологии математического моделирования и обработки данных». В части 3 сборника представлены материалы по направлению «Технология, оборудование и автоматизация машиностроительных производств. Материаловедение, технологические процессы и аппараты». В части 4 сборника представлены материалы по направлению «Энергетика». В части 5 сборника представлены материалы по направлению «Электротехника, электромеханика и электротехнологии». В части 6 сборника представлены материалы по направлению «Электроника и биомедицинская техника». В части 7 сборника представлены материалы по направлению «Экономика и управление». В частях 8 и 9 сборника представлены материалы по направлению «Гуманитарные науки и современность». В части 10 сборника представлены материалы по направлению «Актуальные проблемы авиаракетостроения». В части 11 сборника представлены материалы по направлению «Экология и техносферная безопасность».

---

---

УДК 62(063)  
ББК 72.5я431

ISBN 978-5-7782-5130-4 (Ч.2)  
ISBN 978-5-7782-5128-1

© Коллектив авторов, 2024  
© Новосибирский государственный  
технический университет, 2024

**Секция ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ**  
**ФИЗИЧЕСКИХ ПРОЦЕССОВ В ТЕХНОЛОГИЯХ**  
**И ПРИРОДНЫХ ЯВЛЕНИЯХ**

**РАЗРАБОТКА ВЫЧИСЛИТЕЛЬНЫХ СХЕМ РЕШЕНИЯ  
ОСЕСИММЕТРИЧНЫХ ЗАДАЧ С ИСПОЛЬЗОВАНИЕМ  
СКАЛЯРНЫХ ГРАНИЧНЫХ ЭЛЕМЕНТОВ**

**С.Д. Авдеева**

Новосибирский государственный технический университет,  
г. Новосибирск, avdeeva199915@gmail.com  
Научный руководитель: Рояк М.Э., д.т.н., профессор

*Данная работа посвящена разработке вычислительных схем для решения осесимметричных задач методом граничных элементов. Для вычисления осесимметричных задач магнитостатики реализован подход с совместным использованием скалярных конечных и граничных элементов. Проведено исследование на вычисление сил с использованием двух различных постановок. Для обеих постановок проведено сравнение числа неизвестных, размеров матриц в памяти, времени генерации матриц и времени решения СЛАУ.*

*This paper is about the development of numerical schemes for solving axisymmetric problems with the boundary element method. For the solution of axisymmetric magnetostatics problems, an approach with the joint use of scalar finite and boundary elements has been implemented. The research on the forces computation using two different formulations has been performed. The number of unknowns, matrix sizes in memory, time of matrix generation and SLAE solution have been compared for both formulations.*

Одним из эффективных методов решения краевых задач, описываемых дифференциальными уравнениями в частных производных, является метод граничных элементов (МГЭ). Он основан на представлении решения в виде интегралов по границе области.

Чтобы получить фундаментальное решение для осесимметричной области, необходимо перейти в фундаментальном решении для трехмерной области к цилиндрическим координатам и проинтегрировать его от 0 до  $2\pi$  по переменной  $\varphi$ . Выражения для фундаментального решения и его производной по нормали имеют вид

$$G(R, \hat{R}) = \int_0^{2\pi} \frac{1}{4\pi} \frac{1}{|R - \hat{R}|} d\varphi = \frac{K(m)}{\pi\sqrt{a+b}}, \quad (1)$$

$$\frac{\partial G(R, \hat{R})}{\partial \hat{n}_R} = \frac{1}{\pi\sqrt{a+b}} \left[ \frac{n_r [E(m) - K(m)]}{2r} - \frac{n_r(r - \hat{r}) + n_z(z - \hat{z})}{(a-b)} E(m) \right], \quad (2)$$

где  $K(m)$  – полный эллиптический интеграл I рода,  $E(m)$  – полный эллиптический интеграл II рода,  $a = r^2 + \hat{r}^2 + (z - \hat{z})^2$ ,  $b = 2r\hat{r}$ ,  
 $m^2 = \frac{4r\hat{r}}{(r + \hat{r})^2 + (z - \hat{z})^2}$ .

Одним из основных недостатков МГЭ является то, что его возможно применять только для однородных подобластей, в которых параметры среды являются константными функциями. Данный недостаток можно обойти путем применения подхода с совместным использованием метода конечных (МКЭ) и граничных элементов.

Постановка с использованием двух скалярных потенциалов подразумевает разделение всей расчетной области на две подобласти. В подобласти полного потенциала есть ферромагнетика и нет сторонних токов, а в подобласти неполного потенциала, наоборот, нет ферромагнетиков и есть сторонние токи. При этом в области с неполным потенциалом магнитная проницаемость равна 1. Напряженности магнитного поля в каждой из подобластей имеют вид

$$\vec{H} = \begin{cases} \vec{H}^U = \vec{H}^c - \text{grad}U & \text{в подобласти неполного потенциала,} \\ \vec{H}^P = -\text{grad}P & \text{в подобласти полного потенциала,} \end{cases} \quad (3)$$

где  $\vec{H}^c$  – магнитное поле, создаваемое токами в вакууме.

Уравнение для нахождения полного потенциала имеет вид

$$-\text{div}(\mu \text{grad}P) = 0. \quad (4)$$

Уравнение для нахождения неполного потенциала имеет вид

$$-\text{div}(\text{grad}U) = 0. \quad (5)$$

Поскольку в уравнении для полного потенциала присутствует магнитная проницаемость, которая в общем случае не является константой, то в подобласти полного потенциала будет применяться МКЭ, а в подобласти неполного потенциала – МГЭ.

Модель для тестирования реализованной постановки представлена на рисунке 1. Данная модель использовалась для расчета сил в статье [1]. Модель состоит из двух стальных цилиндров, находящихся на расстоянии 5 мм друг от друга. На нижнем цилиндре находится катушка с 291 витком.



Рис. 1 – Модель для расчета сил

В программном комплексе Tetna была построена модель и были выданы значения  $r$  компоненты напряженности поля. Для проверки правильности полученных значений была решена аналогичная задача в трехмерных координатах. В таблице 1 представлены полученные результаты в некоторых точках для двумерной и трехмерной задач.

Таблица 1 – Значения напряженности поля

Точка		Двумерная задача	Трехмерная задача
$r$	$z$		
0.004	0.115	3.9085	3.9782
0.012	0.055	6.4579	6.4696
0.035	0.08	13099	13095

Также была вычислена сила, действующая на верхний цилиндр, для данной постановки и двумерной постановки с вектор-потенциалом. В таблице 2 представлены результаты для линейной задачи на дважды удвоенной сетки. В первом столбце указана сила тока в катушке, во втором – экспериментальные данные, в третьем – сила для постановки с вектор-потенциалом, а в последнем – для двух скалярных потенциалов.

Таблица 2 – Значения напряженности поля

Сила I, A	$F_n$ , Н	$F_{vec}$ , Н	$F_{scal}$ , Н
6	1.4710	1.4997	1.5285
7	2.0594	2.0412	2.0705
8	2.7459	2.6661	2.7174
10	4.4130	4.1657	4.2460
12	6.1782	5.9986	6.1142

Значения вычисляются достаточно близко к экспериментальным, но не совпадают с ними, вероятно, из-за решения линейной задачи.

В таблицах 3-6 представлено сравнение числа неизвестных, размеров матриц в памяти, времени генерации матриц и времени решения СЛАУ для двух постановок. Во втором и третьем столбцах указаны значения в совместной постановке для МГЭ и МКЭ соответственно, а в последнем – для МКЭ в постановке с вектор-потенциалом.

**Таблица 3 – Количество элементов в сетке**

Уровень вложенности сетки	$scal_z$	$scal_k$	$scal_z + scal_k$	$vec$
0	92	1 040	1 132	2 998
1	184	4 160	4 334	11 992
2	368	16 640	17 008	47 968

**Таблица 4 – Размеры матриц в памяти (мб)**

Уровень вложенности сетки	$scal_z$	$scal_k$	$scal_z + scal_k$	$vec$
0	0.067	0.062	0.129	0.245
1	0.264	0.274	0.538	0.971
2	1.045	0.935	1.980	3.863

**Таблица 5 – Время генерации матриц (с)**

Уровень вложенности сетки	$scal_z$	$scal_k$	$scal_z + scal_k$	$vec$
0	0.967	0.036	1.003	0.085
1	3.653	0.131	3.784	0.328
2	14.08	0.444	14.52	1.250

**Таблица 6 – Время решения СЛАУ (с)**

Уровень вложенности сетки	$scal_z + scal_k$	$vec$
0	0.0161	0.0181
1	0.0582	0.0764
2	0.2184	0.3232

Таким образом, использование МГЭ в непроводящей подобласти позволяет уменьшить число неизвестных. Однако время решения линейной задачи в постановке с двумя скалярными потенциалами намного больше за счет генерации матриц для МГЭ. Для обхода этого недостатка существуют различные подходы к ускорению данного метода. При решении нелинейной задачи расхождение в генерации СЛАУ будет нивелировано, начиная примерно с 10 итераций.

Литература:

1. *Рояк М. Э.* Расчет силы магнитного поля в магнитных системах со стационарным током = The magnetic field force calculation in magnetic systems with stationary current / М. Э. Рояк, Т. Р. Дербышева, С. Д. Авдеева [и др.]. – DOI 10.21883/PJTF.2022.19.53595.19337. – Текст: непосредственный // Письма в Журнал технической физики = Pis'ma v Zhurnal tekhnicheskoi fiziki. – 2022. – Т. 48, № 19. – С. 39–42.
2. *Бенерджи П., Баттерфилд Р.* Методы граничных элементов в прикладных науках: Пер. с англ. – М.: Мир, 1984. – 494 с.
3. *Бреббия К., Теллес Ж., Вроубел Л.* Методы граничных элементов: Пер. с англ. – М.: Мир, 1987. – 524 с.
4. *Вержбицкий В. М.* Основы численных методов: Учебник для вузов. – М.: Высш. шк., 2002. – 840 с.
5. *Соловейчик Ю. Г., Рояк М. Э., Персова М.Г.* Метод конечных элементов для решения скалярных и векторных задач: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2007. – 896 с.
6. *Ступаков И. М.* Разработка алгоритмов решения задач магнитостатики с использованием метода граничных элементов: дис. канд. техн. наук: 05.13.18 / Ступаков Илья Михайлович. – Новосибирск, 2016.
7. *Abramowitz M., Stegun I. A.* Handbook of mathematical functions. – New York: Dover, 1970. – 1046 p.
8. *Brebbia C. A., Walker S.* Boundary element techniques in engineering. – Boston: Newnes-Butterworths, 1980. – 210 p.
9. *Singh J., Gliere A., Achard J.-L.* A multipole expansion-based boundary element method for axisymmetric potential problem // Engineering analysis with boundary elements. – 2009. – Vol. 33, no. 5. – P. 654-660. – DOI: 10.1016/j.enganabound.2008.10.002.
10. *Steinbach O.* Numerical approximation methods for elliptic boundary value problems. – New York: Springer Science, 2008. – 386 p.



# РАЗРАБОТКА АЛГОРИТМА ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЯ ПРИ РЕНТГЕННОГРАФИИ ОБЪЕКТОВ ВЫСОКОЙ ПЛОТНОСТИ

А.А.Зотов

Новосибирский государственный технический университет,

г. Новосибирск, kalodor@mail.ru

Научный руководитель: Григорьев Д.Н., к.ф.-м.н., доцент

*В настоящее время ведется разработка нового детектора на основе сцинтилляционных кристаллов для проведения рентгенографических исследований объектов с высокой оптической плотностью. Для регистрации света в них используются многоэлементные КМОП-фотоприемники. Так как на одно волокно приходится примерно 4500 пикселей и положение оптических волокон относительно пикселей фотоприемника задается конструкцией крепления недостаточно точно, то необходимо разработать алгоритм для восстановления изображений, получаемых с многоэлементных фотоприемников. В данной работе представлен алгоритм, разработанный для анализа данных, описана логика его работы и полученные результаты.*

*Currently, a new detector based on scintillation crystals is being developed for X-ray studies of objects with high optical density. To register light in them, multi-element CMOS photodetectors are used. Since there are approximately 4,500 pixels per fiber and the position of the optical fibers relative to the photodetector pixels is not set precisely enough by the mounting tool, it is necessary to develop an algorithm for restoring images obtained from multi-element photodetectors. This paper presents an algorithm developed for data analysis, describes the logic of its operation and the results obtained*

Рентгенография – это исследование внутренней структуры объекта путем просвечивания его рентгеновскими лучами. На Рисунке 1 представлена схема проведения рентгенографических исследований [1].

Традиционно для проведения рентгенографических исследований используется метод, в котором свет от сцинтилляционного конвертора передается на матричный фотоприемник с помощью объективов. Основной недостаток такой схемы, малая величина светосбора, которая составляет не более 0.2. Замена традиционной оптической системы на волоконную оптику позволила увеличить величину светосбора в несколько десятков раз.

Разработанный детектор для проведения рентгенографических исследований состоит из сцинтилляционного конвертора, волоконной оптики и многоэлементных фотоприемников. Свет, возникший в результате взаимодействия гамма-квантов с веществом сцинтилляционного конвертора, передается по волоконной оптике к многоэлементным фотоприемникам. Положение оптических волокон относительно пикселей фотопри-

емника задается конструкцией крепления недостаточно точно. Количество пикселей, соответствующее одному волокну на изображении, равно примерно 4600 штук. Поэтому была поставлена задача: разработать автоматизированный алгоритм обработки данных с фотоприемников. Созданный автоматизированный алгоритм повысил скорость обработки информации на несколько порядков по сравнению с ручной обработкой сигналов.

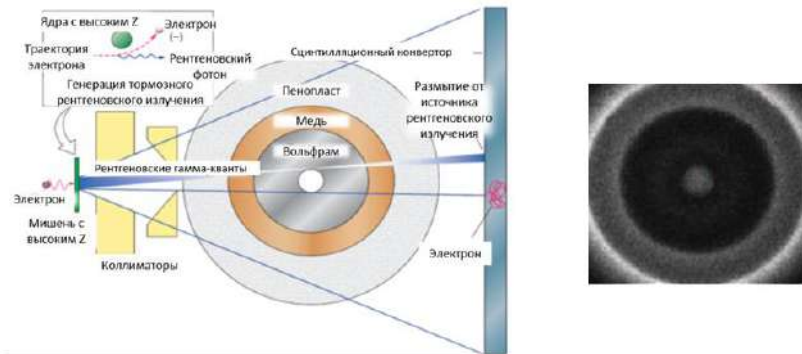


Рис. 1 –Схема проведения рентгенографических исследований(слева), рентгенограмма исследуемого объекта(справа).

Логика работы алгоритма для восстановления изображений состоит из нескольких этапов. На первом этапе из снимков, полученных с фотоприемников, сначала вычитается уровень нулевого сигнала. На втором этапе проводится цифровая обработка снимков, выделяющая границы волокна. На третьем этапе границы сохраняются и составляется карта волокон: каждому волокну присваивается уникальный индекс. Четвертый этап заключается в том, что сигналы с волокон, которые приходят на один сцинтилляционный кристалл суммируются и таким образом определяется величина сигнала в кристалле. Совокупность величин сигналов в кристаллах создает рентгенограмму исходного изображения.

Для проверки работы алгоритма и измерения пространственного разрешения макета детектора был проведен эксперимент с использованием вольфрамового сегмента. Была проведена серия измерений при перемещении сегмента в горизонтальном направлении поперёк потока с шагом 0,2 мм, как показано красной стрелкой на Рисунке 2.

Полученные в результате работы алгоритма, рентгенограммы аппроксимировались функцией  $erf(x)$  (Рисунок 3).

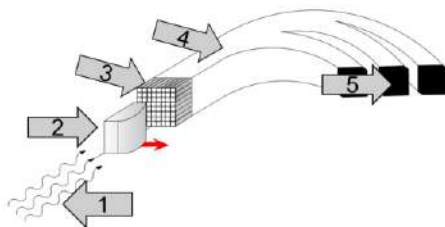


Рис. 2 – Схематическое построение эксперимента по измерению пространственного разрешения макета детектора: 1 – поток гамма-квантов; 2 – вольфрамовый сегмент; 3 – сцинтилляционный конвертор; 4 – волоконная оптика; 5 – многоэлементные фотоприемники.

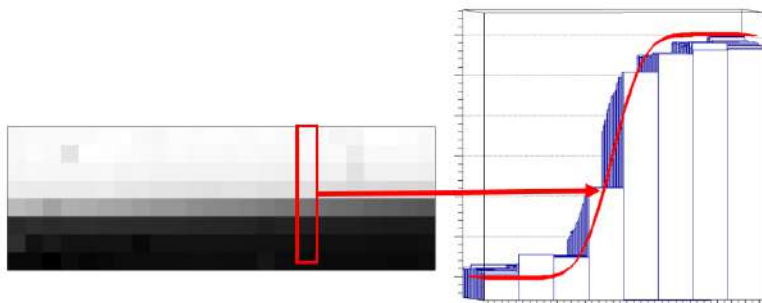


Рис. 3 – Рентгенограмма вольфрамового сегмента (слева), Аппроксимация рентгенограммы с помощью функции  $\text{erf}(x)$  (справа).

Измеренное размытие границы сегмента составило  $1,05 \pm 0,03$  мм, что согласуется с численным моделированием физических процессов взаимодействия  $\gamma$ -квантов с веществом.

На Рисунке 4 представлена зависимость восстановленного положения границы от смещения сегмента, аппроксимированная линейной функцией. Пространственное разрешение оценивалось по распределению разностей восстановленной координаты и смещения сегмента и составило менее 0,04 мм.

В ходе работы был разработан автоматизированный алгоритм восстановления сигналов с изображений, получаемых с многоэлементных фотоприемников при рентгенографических исследованиях объектов. Данный алгоритм применен к данным, полученным в результате эксперимента по измерению пространственного разрешения макета сцинтилляционного детектора с высоким разрешением. Полученные результаты согласуются с моделированием физических процессов взаимодействия  $\gamma$ -квантов с веществом.

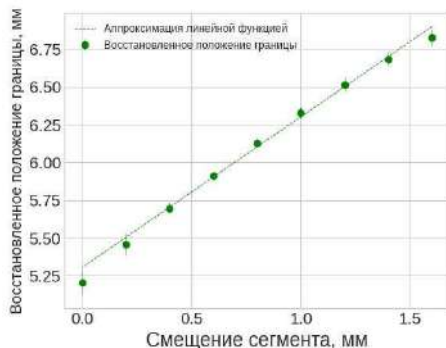


Рис. 4 – Зависимость восстановленного положения границы от смещения сегмента.

#### Литература:

1. Flash Radiography. Historical Origins. Cunningham G. S., Morris C. // Los Alamos Science. – 2003. – nr 28. – P. 76–91. – URL: <https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-03-5305>. – Дата обращения: 25.05.2023.

2. X-ray detectors for digital radiography / M. J. Yaffe, J. A. Rowlands. – Текст : электронный // Physics in Medicine and Biology. – 1997. – Vol. 42, nr 1. – P. 1–39. – URL: <https://doi.org/10.1088/0031-9155/42/1/001>. – Дата публикации: 16.08.1997.

3. Multi-MeV Flash Radiography in Shock Physics Experiments: Specific Assemblages of Monolithic Scintillating Crystals for Use in CCD-Based Imagers / D. Partouche-Sebban, I. Abraham, S. Lauriot, C. Missault. – Текст : электронный // X-Ray Optics and Instrumentation. – 2010. – Vol. 2010. – P. 1–9. – URL: <https://doi.org/10.1155/2010/156984>. – Дата публикации: 17.01.2010.

# ПРОБЛЕМЫ РЕШЕНИЯ ЗАДАЧ СТЕФАНА РАЗРЫВНЫМ МЕТОДОМ ГАЛЁРКИНА С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНОГО ПАКЕТА GMSH

А.В. Кислин

Новосибирский государственный технический университет,  
г. Новосибирск, [kislin.art@yandex.ru](mailto:kislin.art@yandex.ru)

Научный руководитель: Иткина Н.Б., к.т.н., доцент

*В данной статье рассматривается решение двух основных проблем, возникающих при моделировании процессов фазового перехода с помощью метода разрывного Галёркина в сравнении со стандартным методом конечных элементов: построение неконформной конечноэлементной сетки и интегрирование базисных функций по межэлементным граням.*

*This article discusses the solution to two main problems that arise when modeling phase transition processes using the discontinuous Galerkin method in comparison with the standard finite element method: constructing a non-conformal finite element mesh and integrating basis functions along interelement faces.*

Задача Стефана – это особый вид краевой задачи, описывающий процесс фазового перехода различных веществ при заданных условиях. Её характерной отличительной чертой является наличие дополнительных условий на движущейся границе раздела двух фаз, называемой фронтом.

Задача Стефана позволяет моделировать процессы таянья и кристаллизации веществ. Соответственно такие математические модели могут быть использованы при решении задач строительства, добычи газовых гидратов и экологии (вечная мерзлота и т.д.).

Для решения задачи Стефана зачастую используют конформный метод конечных элементов. Однако у неё есть важная особенность. В качестве её решения ищется поле температуры, которое вблизи фронта имеет достаточно резкое изменение. Поэтому рядом с линией раздела двух фаз необходимо произвести достаточно мелкое разбиение конечноэлементной сетки. А поскольку конформный метод конечных элементов требует конформности сетки, то её дробление в некоторой части, неизбежно повлечёт за собой дробление всей расчётной области, что очевидно приведёт к увеличению затрат по памяти и времени. Поэтому для решения задачи с движущейся границей эффективнее будет использовать неконформные конечноэлементные методы. Одним из таких методов является метод разрывного Галёркина.

В отличие от конформного метода конечных элементов в методе разрывного Галёркина в качестве базисного пространства можно выбрать

пространство разрывных функций, что позволяет искать решение краевой задачи на неконформной сетке. Но, с другой стороны, это серьёзно усложняет вариационную постановку, к которой помимо интегрирования по самим конечным элементам, добавляется интегрирование по их межэлементным граням внутри области. В дополнение к этому остро встаёт вопрос о самом способе построения сетки, в которой нет обязательного условия полного сопоставления границ конечных элементов.

Резюмируя вышесказанное, можно выделить следующие основные проблемы метода разрывного Галёркина в сравнении с конформным методом конечных элементов:

1. Построение неконформной конечноэлементной сетки.
2. Интегрирование базисных функций на межэлементных гранях внутри области.

Поскольку построение конечноэлементной сетки сама по себе задача достаточно трудоёмкая, в отрыве от реализации самого метода решения краевой задачи, то оптимальным решением будет использовать для этой цели готовый программный пакет. В качестве данного пакета был выбран свободно-распространяемый построитель тетраэдральных сеток GMSH. Его главным и решающим преимуществом перед остальными пакетами является возможность его бесплатного использования без ущерба его функциональности. Однако у данного пакета есть в нашем случае один существенный недостаток - он способен строить только конформные сетки. Но всё же использовать его для построения расчётной области с неконформным разбиением возможно.

У GMSH есть два основных модуля: модуль модели и модуль построителя сетки. В модуле модели задаётся описание расчётной области: указываются базовые точки, которые соединяются в линии, а затем в объёмные фигуры. В модуле сетки, полученные из модели линии и фигуры уже разбиваются на конечные элементы.

Чтобы в GMSH построить неконформную сетку, можно расчётную область разбить на несколько подобластей в зависимости от их разбиений. Важно, чтобы при пересечении этих областей их общие точки, на которых они базируются в модели, идентифицировались в ядре GMSH по-разному в зависимости от того, через какую подобласть к ним обращаются. Добиться этого можно двумя способами. Первый – в скрипт-файле GMSH напрямую создать несколько точек с одинаковым геометрическим расположением, но разными идентификаторами (tag-ами). Второй – в графическом интерфейсе намеренно создать подобласти чуть в отдалении друг от друга, а потом с помощью операции перемещения их соединить, но обязательно предварительно указав параметры разбиения

каждой подобласти. Для обоих из этих способов так же необходимо сделать указание GMSH на запрет удаления дубликатов получаемых точек, иначе на выходе можно получить расчётную область с конформным разбиением.

К сожалению, хоть данные способы и позволяют построить неконформную сетку, но они не решают проблемы с определением межэлементных границ, вернее решают её не полностью. GMSH способен выдать межэлементные границы внутри конформных подобластей, но не на их границах. Чтобы получить межэлементные грани на плоскостях с неконформной стыковкой конечных элементов, необходимо предварительно указать поверхностям двух подобластей, которые будут пересекаться, одну физическую группу. Затем, с её помощью взяв информацию по разбиению этих поверхностей на треугольники (на поверхности области очевидно будут грани конечных элементов, а грань тетраэдра – треугольник), необходимо их явно добавить в модуль модели, после чего применить к ним операцию фрагментирования, встроенную в GMSH. В качестве объекта для этой операции необходимо указать треугольники с поверхности одной подобласти, а в качестве инструмента треугольники с поверхности другой. В итоге мы получим новое треугольное разбиение поверхности пересечения подобластей, где каждый треугольник полностью принадлежит не более чем двум конечным элементам. Для быстрого определения к каким конечным элементам принадлежит тот или иной полученный треугольник можно реализовать и воспользоваться алгоритмом геометрического поиска с помощью квадродерева.

Поскольку в данном случае все межэлементные границы будут сведены к треугольникам, то и проблема интегрирования базисных функций будет сведена к интегрированию функций на треугольниках. Отметим, что из-за неконформности сетки, полученные межэлементные границы могут быть только частью граней двух элементов, аналитический расчёт интегралов на них в общем случае не предоставляется возможным. Поэтому для интегрирования будем использовать квадратурные формулы Гаусса.

Рассмотрим произвольный треугольник с координатами  $(x_1, y_1, z_1)$ ,  $(x_2, y_2, z_2)$ ,  $(x_3, y_3, z_3)$ . Поскольку треугольник это плоская фигура, отобразим его на одну из координатных плоскостей, для определённости выразив координату  $z$  через  $x$  и  $y$ , и в соответствии с этим заменим интегрируемую функцию  $F(x, y, z(x, y))$  на  $f(x, y)$ , переходя от интеграла в трёхмерном пространстве к интегралу в двумерном:

$$f(x, y) = F(x, y, z(x, y)) * \sqrt{(1 + a_1^2 + a_2^2)} \quad (1)$$

где  $a_1$  и  $a_2$  – коэффициенты проецирования плоскости треугольника на плоскость XY, которые находятся из следующей системы уравнений:

$$\begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \quad (2)$$

Теперь произведём замену переменных, которая переведёт произвольный треугольник с точками  $(x_1, y_1)$ ,  $(x_2, y_2)$  и  $(x_3, y_3)$  в единичный.

$$\begin{aligned} x_3 + (x_1 - x_3)\xi_1 + (x_2 - x_3)\xi_2 &= x \\ y_3 + (y_1 - y_3)\xi_1 + (y_2 - y_3)\xi_2 &= y \end{aligned} \quad (3)$$

где  $\xi_1, \xi_2$  изменяются в промежутке  $[0, 1]$ , причём  $\xi_1 + \xi_2 \leq 1$ .

Поскольку метод Гаусса работает только на прямоугольниках то необходимо произвести ещё одну замену, которая переведёт полученный единичный треугольник в единичный квадрат:

$$\xi_1 = lm \quad \xi_2 = m(1 - l) \quad (4)$$

Не стоит забывать, что с каждой заменой переменных необходимо менять и подынтегральную функцию, умножая её на Якобиан перехода. Таким образом с учётом всех замен можно перейти от интегрирования по треугольнику в трёхмерном пространстве к интегрированию по квадрату в двумерном, где уже можно будет воспользоваться квадратурной формулой Гаусса.

По итогу с помощью программного пакета GMSH и некоторых интегральных аналитических выкладок можно решить две основные проблемы метода Разрывного Галёркина, связанные с построением сетки и интегрированием по межэлементным граням. А уже само применение данного метода позволит значительно уменьшить количество конечных элементов при решении задачи Стефана, что может существенно сократить затраты по памяти и времени при моделировании процессов фазового перехода.

Литература:

1. Arnold D.N., Brezzi F., Cocburn B., Marini D. Unified analysis of discontinuous Galerkin methods for elliptic problems // SIAM J. Numer. Anal.—2002. — Vol. 39, N 5. — P. 1749–1779.
2. GMSH tutorial [Электронный ресурс] // URL. <https://web.archive.org/web/20231109160707/https://gmsh.info/doc/texinfo/gmsh.html#Gmsh-tutorial> (дата обращения 15.11.2023).
3. Gazeta de Souza L. H., Aparecido J. B. Numerical integration by Gauss-Legendere quadrature over triangular domains [Электронный ресурс] // URL.



<https://www.abcm.org.br/anais/cobem/2007/pdf/COBEM2007-1614.pdf>  
(дата обращения 03.10.2023).

4. Глушков С.А. Применение стабилизированных конформных и неконформных конечноэлементных схем к решению задач тепло- и массообмена: магистерская диссертация: 01.03.02. – Новосибирск: НГТУ, 2019. – 116 с.

5. Соловейчик Ю.Г., Рояк М.Э., Персова М.Г. Метод конечных элементов для решения скалярных и векторных задач // Новосибирск: НГТУ, 2007. – 896 с.

6. Кузнецов Ф.А., Истомин В.А., Родионова Т.В. Газовые гидраты: исторический экскурс, современное состояние, перспективы исследований // Российский Химический журнал. – 2003 – Т. XLVII № 3 – С. 5 – 18.

## **ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ПРИ ПРОЕКТИРОВАНИИ УСТРОЙСТВ ЗАЩИТЫ ОПЕРАТОРА ДОРОЖНО-СТРОИТЕЛЬНОЙ ТЕХНИКИ ОТ ПАДАЮЩИХ ПРЕДМЕТОВ**

**С.А. Кислицын<sup>1,2</sup>, И.И. Баранов<sup>1,2</sup>, И.Д. Шадрин<sup>1,2</sup>**

<sup>1</sup>Уральское конструкторское бюро транспортного машиностроения,  
г. Нижний Тагил, [mail@ukbtm.uvz.ru](mailto:mail@ukbtm.uvz.ru);

<sup>2</sup>Нижнетагильский технологический институт (филиал) Уральского  
Федерального университета имени первого Президента России  
Б.Н. Ельцина, г. Нижний Тагил

**Научный руководитель: Хмельников Е.А.<sup>2</sup>, член-корреспон-  
дент РАН, д.т.н., профессор, заведующий кафедрой «Специаль-  
ное машиностроение»**

*В статье представлен процесс создания имитационной модели взаимодействия системы «ударник-защищаемый объект». Проведено сопоставление результатов моделирования с результатами экспериментального исследования рассматриваемого процесса. В результате чего подтверждена возможность проводить виртуальные эксперименты на верифицированных моделях при проектировании защитных устройств.*

*This article presents the process of creating a simulation model interaction the "striker - protected object" system. The simulation results were compared with the results of an experimental study. As a result, the possibility of conducting virtual experiments on verified models in the design of protective devices was confirmed.*

При эксплуатации дорожно-строительной техники на таких операциях, как содержание автомобильных дорог, благоустройство участков, коммунальных работах, прочих работах на строительных объектах, а также при лесоразработках, основную опасность для оператора представляет падение на кабину различных предметов: кирпичей, небольших бетонных блоков, обломков породы, деревьев, ручного инструмента и тому подобное. Для обеспечения защиты оператора от падающих предметов используется система конструктивных элементов, смонтированных на машине, называемая *FOPS* (*Falling-object protective structures* – Устройства защиты от падающих предметов) [1].

При сертификации изделий дорожно-строительной техники требуется обязательное проведение испытаний системы *FOPS*. Согласно [1] и [2], защитные свойства *FOPS* должны обеспечивать падение груза с высоты, достаточной для развития энергии 11,6 кДж. При этом проникновение элементов кабины или системы *FOPS* в зону ограничения предельно допустимой деформации не допускается.

В соответствии с [1], зона ограничения предельно допустимой деформации представляет собой защищённый лимитированный объём *DLV* (*deflection-limiting volume*): фигура с прямоугольными очертаниями, по размерам приблизительно соответствующая антропометрическим данным оператора (мужчина высокого роста в положении сидя в обычной рабочей одежде) [3].

При проектировании новых изделий дорожно-строительной и лесозаготовительной техники уделяется большое внимание безопасной работе оператора. Поэтому при разработке нового изделия очень важно спроектировать максимально безопасную, но в то же время наиболее лёгкую и технологичную систему защиты оператора.

Проведение испытаний всех возможных конструктивных решений схем защиты повлекло бы за собой значительное увеличение времени проектирования и существенное удорожание готового изделия. Поэтому при разработке новых изделий широкое применение нашли методы математического моделирования, использование которых позволяет натурные испытания проводить лишь для подтверждения оптимальности принятых решений.

В данной статье рассматривается процесс моделирования и испытания кабины универсального малогабаритного погрузчика, оборудованной системой *FOPS*, проведено сравнение результатов моделирования и эксперимента.

Перед проведением виртуального и натурального экспериментов был выбран груз призматической формы с размерами основания  $0,5 \times 0,5$  м массой

473 кг (исходя из наличия в производстве). Для развития энергии взаимодействия 11,6 кДж высота падения данного груза должна составлять 2,5 м.

Для проведения моделирования падения груза на кабину создана конечно-элементная модель кабины и груза. Для описания поведения материала груза, кабины и элементов FOPS была выбрана модель упругопластического материала. Моделирование выполнялось в Лагранжевой постановке. Грузу задана скорость взаимодействия с кабиной, соответствующая кинетической энергии, приобретаемой грузом массой 473 кг при падении с высоты 2,5 м.

Расчётная схема и результаты визуализации остаточных деформаций кабины и элементов системы FOPS после падения груза представлены на рисунке 1.

В результате моделирования определено, что разрушения элементов кабины и системы FOPS не происходит.

Результаты расчётной вертикальной составляющей остаточной деформации крыши в определённых точках занесены в таблицу 1.

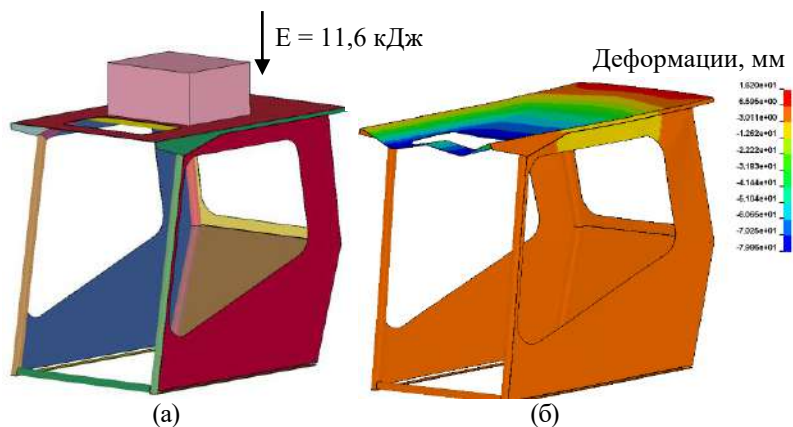


Рис. 1 – Расчётная схема (а) и максимальные деформации (б)

Для оценки адекватности расчётной модели и методики расчёта был проведён натурный эксперимент, в ходе которого был произведён сброс технологического груза с ковша экскаватора по центру крыши кабины погрузчика.

Положение груза и кабины перед началом эксперимента представлено на рисунке 2.



Рис. 2 – Положение груза и состояния кабины в начале эксперимента

После испытаний была определена вертикальная составляющая остаточной деформации крыши в заданных точках (см. таблицу 1), проведён осмотр каркаса кабины, силовых элементов системы FOPS и сварочных швов – разрушений не обнаружено.

Деформация каркаса кабины после взаимодействия с грузом представлена на рисунке 3.

Расчётные и экспериментальные значения вертикальной составляющей остаточной деформации крыши представлены в таблице 1.



Рис. 3 – Деформация кабины после падения груза

**Таблица 1 – Расчётные и экспериментальные значения вертикальной составляющей остаточной деформации крыши**

Номер точки	Величина деформации крыши в точке, мм									
	1	2	3	4	5	6	7	8	9	10
Расчётные значения	30	52	49	44	34	50	67	80	70	58
Экспериментальные значения	31	50	48	43	35	50	69	81	69	57

## Выводы

В результате воздействия падающего груза с энергией 11,6 кДж на кабину разрушений силовых элементов и сварочных швов не произошло.

Устройство защиты FOPS испытания выдержало, так как в результате деформации конструкции нет проникновения её элементов в объём ограничения деформации (расстояние от потолка кабины до головы оператора высокого роста составило приблизительно 200 мм).

Результаты расчёта имеют высокую сходимость с результатами эксперимента, что позволяет применять данную методику при проектировании новых изделий, тем самым значительно сокращая число натурных испытаний.

## Литература:

1. ГОСТ Р ИСО 3449-2009 Машины землеройные. устройства защиты от падающих предметов. лабораторные испытания и технические требования. М. : Стандартинформ, 2010. 11 с.
2. ГОСТ ISO 8083-2011 Машины для леса. устройства защиты от падающих предметов. технические требования и методы испытаний. М. : Стандартинформ, 2020. 10 с.
3. ГОСТ Р ИСО 3411-2011 Машины землеройные. Антропометрические данные операторов и минимальное рабочее пространство вокруг оператора. М. : Стандартинформ, 2012. 9 с.

## **ПОСТРОЕНИЕ КОНЕЧНОЭЛЕМЕНТНЫХ СЕТОК ПРИ МОДЕЛИРОВАНИИ ПРОЦЕССОВ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ С УЧЕТОМ ТРАЕКТОРИИ ГОРИЗОНТАЛЬНЫХ СКВАЖИН**

**Д.А.Леонович**

**Новосибирский государственный технический университет,  
г. Новосибирск, Leonovich.D.A@yandex.ru**

*В работе рассматривается подход к построению конечноэлементных сеток, позволяющий учитывать траекторию скважин. Эта технология особенно важна при моделировании процессов многофазной фильтрации на месторождениях с протяженными горизонтальными скважинами.*

*The paper considers an approach to the construction of finite element meshes, which allows taking into account the trajectory of wells. This technology is especially important when modeling multiphase filtration processes in fields with horizontal wells.*

Использование горизонтальных и наклонных скважин вызывает все больше интереса, так как развитие технологий за последние годы позволило расширить возможности их применения. При разработке углеводородных месторождений горизонтальные скважины позволяют эффективнее применять различные методы повышения нефтеотдачи: например, закачку горячей воды или SAGD (технология парогравитационного дренажа). [1-2] Горизонтальные скважины пробурены вдоль целевого пласта и как правило имеют большую протяженность. Из-за неоднородной структуры геологического слоя при моделировании разработки углеводородных месторождений становится актуальной задача учета траектории этих скважин.

Численное моделирование многофазной фильтрации будет проводиться методом конечных элементов. Для этого на основе закона Дарси для фильтрации жидкостей и газов в пористых средах и закона сохранения массы для каждой из фаз составляется краевая задача для расчета поля давления:

$$-\operatorname{div}\left(\sum_{m=1}^{NP}\frac{\kappa^m}{\mu^m}\mathbf{K}\left(\operatorname{grad}P+(0,0,\rho^m g)^T\right)\right)=f^V(P), \quad (1)$$

с краевыми условиями, определенными на границе  $\partial\Omega=\Gamma_1\cup\Gamma_2$  расчетной области  $\Omega$ ,

$$P|_{\Gamma^1}=P^\Gamma, \quad (2)$$

$$\sum_{m=1}^{NP}\frac{\kappa^m}{\mu^m}\mathbf{K}\left(\operatorname{grad}P+(0,0,\rho^m g)^T\right)\Big|_{\Gamma^2}\cdot\vec{n}=\sum_{m=1}^{NP}f^{m,\Gamma}. \quad (3)$$

$\mathbf{K}$  – тензор структурной проницаемости,  $m$  – номер фазы,  $NP$  – число фаз,  $\kappa^m, \mu^m, \rho^m$  – коэффициент относительной фазовой проницаемости, вязкость и плотность фазы  $m$ , соответственно,  $P$  – давление.  $g$  – ускорение свободного падения,  $f^V$  – объемный источник, возникающий например в результате сжимаемости фаз или из-за других процессов в пласте,  $P^\Gamma$  – давление, заданное на границах  $\Gamma^1$ , а  $\sum_{m=1}^{NP}f^{m,\Gamma}$  – поток смеси сжимаемых фаз, заданный на границах  $\Gamma^2$ .

Задача (1)-(3) решается на каждом временном шаге и на основе найденного поля давления определяются потоки фаз в расчетной области. По ним рассчитываются перетекающие объемы фаз через грани ячеек конечноэлементной сетки и обновляется распределение фаз в модели. Метод численного моделирования более подробно рассмотрен в работах [3].

Скважины в задаче моделирования разработки нефтяных месторождений являются основным источником для поля давления, поэтому вблизи зон перфорации необходима более подробная сетка. Для горизонтальных скважин задаются радиус скважины, траектория ствола, расстояние от начала траектории до начала и конца зон перфорации, коэффициент разрядки (для сгущения сетки вблизи скважины), количество дроблений между скважиной и базовой сеткой, коэффициент, определяющий во сколько раз область встройки скважины должна быть больше радиуса скважины.

Для встройки горизонтальной скважины на основе траектории определяются участки зон перфорации. Базовая сетка ориентирована вдоль осей координат  $OX$ ,  $OY$  и задана в шаблонных координатах по  $z$  (при помощи сплайна по глубине задается рельеф в расчетной области). Поэтому каждая зона перфорации подразбивается на фрагменты по пересечению с линиями базовой сетки и каждый фрагмент встраивается в сетку по прямой, параллельно линиям базовой сетки. Вокруг объекта-скважины задаются дополнительные элементы для сгущения сетки в области зоны перфорации. Чтобы избежать избыточного дробления, используются несогласованные сетки – дополнительные объекты встраиваются только в областях зон перфорации скважин.

Во время моделирования процесса фильтрации элементы, попадающие «внутри» зоны перфорации, считаются непроницаемыми (т.е., их фазовое состояние не меняется), а через грани таких ячеек задан поток смеси в соответствии с режимом работы скважины. Так же для более точного соответствия моделируемому процессу на протяжении всей зоны перфорации задаются специальные краевые условия, позволяющие получить решение по всей длине зоны перфорации скважины с учетом гидростатического давления.

Например, на рисунке 1 зеленым цветом выделена траектория зоны перфорации, синим – линии базовой сетки, белым – элементы сетки, считающиеся «внутри» зоны перфорации.

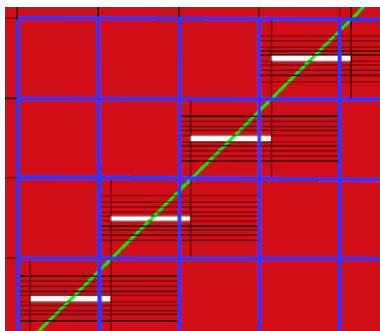


Рис.1 – Фрагмент сетки в плоскости ХУ вблизи с зоной перфорации

Для исследования погрешности аппроксимации за счет поворота фрагментов зоны перфорации (точность которой может регулироваться выбором шага базовой сетки), был проведен следующий численный эксперимент. Модель с двумя горизонтальными параллельными друг другу скважинами была сначала построена так, чтобы скважины были параллельны оси ОУ (назовем это «исходная модель»). После чего модель была повернута на  $45^\circ$  относительно своего центра («модель с поворотом»). Расположение скважин в расчетной области показано зеленым цветом на рисунке 2.

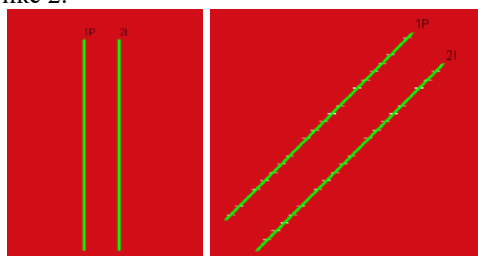


Рис.2 – Схематичное расположение скважин в исходной модели (слева) и после поворота (справа)

Для оценки результатов сравнивались графики суммарного отбора нефти, приведенные на рисунке 3.



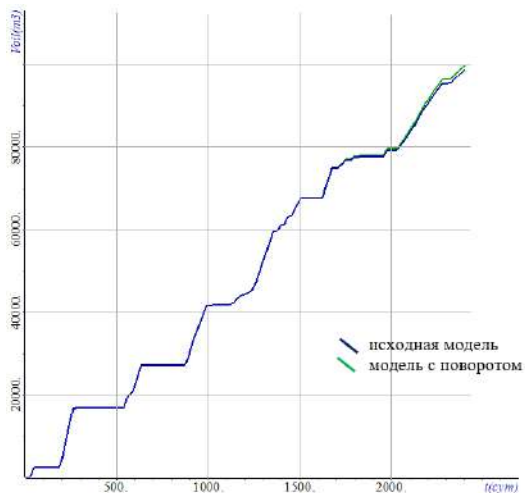


Рис.3 – Графики суммарного отбора нефти для двух моделей

Таким образом, реализованная технология построения сеток позволяет учитывать траектории горизонтальных скважин, что увеличивает точность аппроксимации при моделировании месторождений с несколькими непараллельно расположенными горизонтальными скважинами. Также проведенный численный эксперимент показывает, что точность такой аппроксимации траектории является достаточной: отличие суммарного отбора нефти для модели с поворотом составляет не более 0.2%

#### Литература:

1. Mokheimer, E., Hamdy, M., Abubakar, Z., Shakeel, M. R., Habib, M. A., & Mahmoud, M. A comprehensive review of thermal enhanced oil recovery: Techniques evaluation // *Journal of Energy Resources Technology*, 2019, T. 141, №. 3.
2. Хисамов Р.С., Назимов Н.А., Хайруллин М.Х., [и др.] Оценка профиля притока к стволу горизонтальной скважины по результатам термогидродинамических исследований // *Нефтяное хозяйство*, 2021. №12. С. 114–116.
3. A method of FE modeling multiphase compressible flow in hydrocarbon reservoirs / Y. G. Soloveichik, M. G. Persova, A. M. Grif, A. S. Ovchinnikova, I. I. Patrushev, D. V. Vagin, D. S. Kiselev // *Computer Methods in Applied Mechanics and Engineering*. – 2022. – Vol. 390. – Art. 114468.

# МОДЕЛИРОВАНИЕ ОСАЖДЕНИЯ СЕРЕБРА И ИНДИЯ НА ПОВЕРХНОСТЬ КРЕМНИЯ Si(001)-(2×1) МЕТОДОМ КИНЕТИЧЕСКОГО МОНТЕ-КАРЛО

Д.В. Продан

Сколковский институт науки и технологий, г. Москва,  
dmitrii.prodan@skoltech.ru

Научный руководитель: Юдин Д.И., к.ф.-м.н.

*Для моделирования процессов роста наноструктур на различных подложках широко используется метод кинетического Монте-Карло, позволяющий учитывать стохастическое поведение отдельных атомов на больших пространственных и временных масштабах. В исследовании предлагается модель одновременного осаждения атомов серебра и индия на поверхность кремния.*

*In the modeling of nanostructure formation on various surfaces, the kinetic Monte Carlo method is widely employed. This method allows one to simulate the stochastic behavior of individual atoms with respect to large spatial and temporal scales. In this study, we propose a model for the simultaneous deposition of silver and indium atoms on the reconstructed silicon surface.*

Создание современных материалов и наноустройств плотно связано с изучением технологий выращивания на различных подложках. Развитие подобных технологий сопряжено с разработкой теоретических и численных моделей, способных описывать процессы самоорганизации атомов на поверхности.

Моделирование таких систем осложнено стохастичностью процессов. С одной стороны необходимо учитывать недетерминированное поведение отдельных атомов, с другой стороны – макроскопический размер системы и большие временные масштабы (порядка секунд и больше). Для теоретического описания традиционно используют метод кинетического Монте-Карло (КМК) [1], обладающий высокой вычислительной эффективностью, и позволяющий учесть стохастичность процессов.

В статье рассматривается формирование двумерных структур на поверхности кремния Si(001)-(2×1). Данная реконструированная поверхность образована параллельными рядами димеров кремния и является анизотропной. Кроме того, поверхность кремния может содержать С-дефекты. Для атомов металла на такой поверхности образуется решётка из устойчивых положений, между которыми происходит диффузия. Переход между ними сопровождается преодолением энергетического барьера, профиль которого определяется взаимодействием атома с поверхностью, а, также, при их наличии, с соседними атомами металла и дефектами поверхности. Диффузия атомов металла приводит к образованию островов

– скоплений атомов металла, занимающих соседние устойчивые положения. В качестве центров кластеризации островов могут выступать дефекты кремния.

В рамках предложенного метода на основе КМК рассматривается двумерная решётка, атомы металла могут перемещаться по её разрешённым положениям. При моделировании на каждом шаге происходит либо осаждение нового атома в случайный доступный положение, либо диффузия ранее осаждённого атома, которая может сопровождаться образованием нового острова (присоединение к другому свободному атому или дефекту), присоединением к существующему острову и отсоединением от существующего острова. Диффузия атомов происходит по модифицированной схеме Гиллеспи [2]. Дефекты располагаются случайным образом с заданной плотностью.

Коэффициент вероятности перехода атома через энергетический барьер определяется по формуле Аррениуса (1):

$$p \propto e^{-\frac{E}{kT}}, \quad (1)$$

где  $E$  – величина энергетического барьера – которую можно оценить в рамках теории функционала плотности.

Отдельные положения на решётке могут быть энергетически запрещены для перехода, в таком случае вероятность диффузии или осаждения атомов в эти положения равна нулю.

Ранее другими авторами были разработаны модели осаждения атомов алюминия, галлия и индия [3-5] на поверхность кремния Si(001)-(2×1) при низкой степени покрытия поверхности (до 0,1 монослоя). Для этих металлов характерно формирование островов в виде цепочек, направленных перпендикулярно рядам димеров. Энергетические барьеры при этом зависят от принадлежности атомов островам и конфигурации этих островов.

Более сложная задача – одновременное осаждение атомов двух разных металлов – в настоящее время мало рассмотрена, как с теоретической, так и с экспериментальной точек зрения. В работе [6] представлены изображения полученные после осаждения атомов серебра и индия на поверхность кремния Si(001)-(2×1) полученные в рамках сканирующей туннельной микроскопии (СТМ). Степень покрытия поверхности в результате осаждения составила 0,1 монослоя. Как показано, атомы серебра и индия занимают разные положения на поверхности кремния. Атомы серебра могут образовывать димеры. Авторами работы были выдвинуты предположения по возможным конфигурациям двух металлов с учётом С-дефектов. Отмечается так же характерная перекрёстная форма образующихся структур.

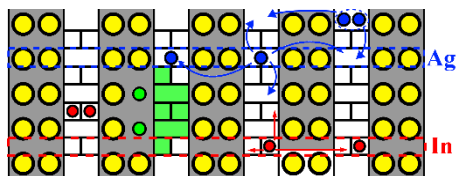


Рис.1 – Модель диффузии атомов серебра (синий) и индия (красный) по поверхности кремния (жёлтый), представленной рядами димеров (отмечены серым фоном). Диффузия серебра и индия возможна по различным положениям. Зеленые кружки – С-дефекты, зеленые области – запрещенные положения.

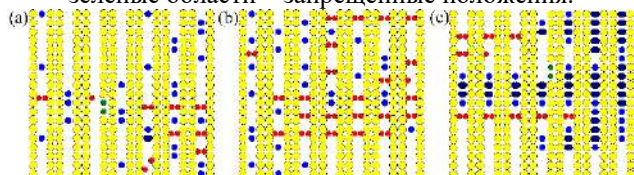


Рис.2 – Фрагменты поверхности с атомами, полученные в рамках развитой методологии при разных температурах: а) 150К, б) 300К и в) 500К. Жёлтым цветом обозначены атомы кремния, красным – индия, синим – серебра. Сдвоенными тёмно-синими кружками обозначены димеры серебра.

На основании упомянутых выше работ и расчётов в рамках теории функционала плотности была разработана модель для системы с одно-временным осаждением серебра и индия. Так как атомы данных металлов диффундируют по разным положениям на решётке, поверхность для биметаллической системы имеет сложную структуру. Так же, в сравнении с монометаллической системой, в предлагаемом подходе существенно расширен перечень конфигураций, в которых может находиться атом металла. Если в монометаллической системе 6 возможных конфигураций, то в биметаллической для двух металлов возможное число конфигураций атомов на поверхности увеличивается до 53 [7]. В силу анизотропии поверхности для каждой конфигурации энергетический барьер рассчитывался вдоль и поперёк рядов димеров.

С использованием модели было проведено исследование качественных и количественных характеристик образующихся структур при разных температурах и на разных этапах осаждения. Полученные структуры сравнивались с экспериментально полученными изображениями. Количественно оценивалась плотность островов, определяемая по формуле (2):

$$N_{isl} = \frac{\theta}{s}, \quad (2)$$

где  $\theta$  – степень покрытия поверхности,  $\bar{s}$  – средний размер островов в количестве атомов.

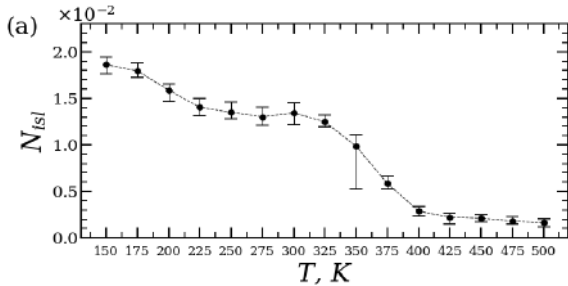


Рис.3 – Плотность островов при разных температурах.

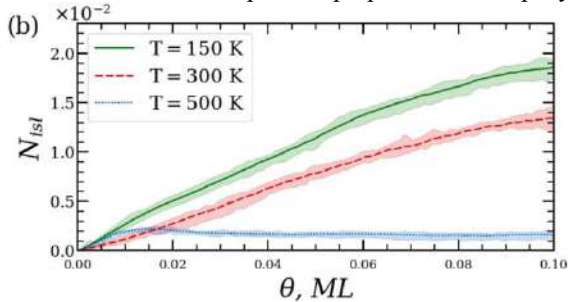


Рис.4 – Динамика плотности островов по мере осаждения при разных температурах.

Качественная картина осаждения при комнатной температуре соответствует перекрёстной структуре, полученной на эксперименте. Отмечено, что при низких температурах образуются небольшие острова размером в несколько атомов, что при увеличении покрытия поверхности приводит к росту плотности островов. С увеличением температуры атомы металлов образуют более крупные острова и плотность островов в ходе осаждения остаётся постоянной.

Литература:

1. Sickafus K.E., Kotomin E.A., Uberuaga B.P. Radiation Effects in Solids. – Springer Dordrecht, 2007. – 592 p.
2. Gillespie, D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions // J. Comp. Phys. – 1976. – Vol. 22, N 4. – P. 403-434. – doi: 10.1016/0021-9991(76)90041-3
3. Structure of low-coverage phases of Al, Ga, and In on Si(100) / Northrup J., Schabel M.C., Karlsson C.J., Uhrberg R. I. G. // Phys. Rev. B. – 1991. – Vol. 44, N 24. – P. 13799-13802. – doi: 10.1103/PhysRevB.44.13799

4. Albao, M.A., Evans, J.W., Chuang, F.C. A kinetic Monte Carlo study on the role of defects and detachment in the formation and growth of In chains on Si(100) // J. Phys. Condens. Matter. – 2009. – Vol. 21, N 20. – P. 405002 – doi: 10.1088/0953-8984/21/40/405002

5. Albia J.R., Albao M.A. Non-Arrhenius temperature dependence of the island density of one-dimensional Al chains on Si(100): A kinetic Monte Carlo study // J. Vac. Sci. Technol. A. – 2015. – Vol. 33, N 2. – P. 021404 – doi: 10.1116/1.4905457

6. Sobotík P., Kocán P., Ošťádal I. Anisotropic alloying: Formation of atomic scale trellis on the Si(100)-(2 × 1) surface // Surf. Sci. – 2018. – Vol. 677. – P. 8-11. – doi: 10.1016/j.susc.2018.05.010

7. An ab initio approach to anisotropic alloying into the Si(001) surface / Prodan D.V., Paradezhenko G.V, Yudin D.I., Pervishko A.A. // Physical Chemistry Chemical Physics. – 2023. – Vol. 25., N. 7 – P. 5501-5509. – doi: 10.1039/d2cp04405f

## **ОПТИМИЗАЦИЯ ВЫЧИСЛЕНИЙ ДЛЯ МОНТЕ-КАРЛО МОДЕЛИ ГЕТЕРОЭПИТАКСИАЛЬНОГО РОСТА Ge НА Si(100)**

**А.Л. Ревун<sup>1,2</sup>, С.А. Рудин<sup>1</sup>, К.В. Павский<sup>1,2</sup>**

**<sup>1</sup>Институт физики полупроводников им. А.В. Ржанова СО РАН,  
г. Новосибирск**

**<sup>2</sup>Сибирский государственный университет телекоммуникаций  
и информатики, г. Новосибирск**

*Рассмотрены методы и алгоритмы моделирования гетероэпитаксиального роста Ge на Si(100) методом Монте-Карло. Проведен анализ вычислительной сложности составляющих алгоритмы блоков и предложено распараллеливание трудоемких циклов с использованием директив открытого стандарта OpenMP.*

*Methods and algorithms for Monte Carlo simulations of heteroepitaxial growth of Ge on Si(100) are considered. The computational complexity of the algorithm blocks is analyzed. Parallelization of time-consuming cycles using directives of the open standard OpenMP is proposed.*

Квантовые точки (КТ) GeSi представляют значительный интерес как с исследовательской так и с практической точки зрения. Совместимость с существующей кремниевой технологией делает их перспективными для создания электронных и электронно-оптических приборов нового поколения [1, 2]. Для изучения роста структур с КТ разработана трехмерная

атомистическая Монте-Карло (МК) модель гетероэпитаксиального роста [3]. Одной из проблем использованной программной реализации модели является высокая трудоемкость вычислений. Например, моделирование осаждения 0.543 нм Ge на подложку Si размерами 27x27x8.15 нм со скоростью 0.00136 нм/с при температуре 500° С (рис. 1) занимает более 1900 ч. При увеличении размера моделируемой структуры и/или температуры существенно возрастает время расчета. Поэтому актуальной задачей является поиск путей оптимизации вычислений, а также разработка параллельных алгоритмов для многоядерных вычислительных систем.

Моделирование роста производится методом Монте-Карло, заключающимся в выполнении последовательности элементарных событий диффузионного прыжка или осаждения атома, выбираемых случайным образом в соответствии с их вероятностями. После каждого элементарного события выполняется «Термический отжиг»: случайным образом выбираются 1% от всех атомов системы, для которых задаются новые случайные смещения вокруг их равновесных положений. Более детальную информацию о физической модели можно найти в [4].

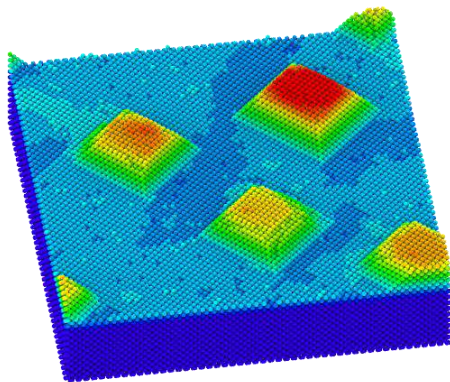


Рис.1 – Результат моделирования осаждения 0.543 нм Ge на подложку Si размерами 27x27x8.15 нм со скоростью 0.00136 нм/с при температуре 500° С. Цвет соответствует высоте рельефа.

На рис. 2 представлена блок-схема «Шаг Монте-Карло» и соответствующие затраты машинного времени на выполнение каждой из его частей. После шага Монте-Карло выполняется «Термический отжиг», который занимает примерно столько же времени. В каждом из выбранных блоков предложено распараллеливание трудоемких циклов с использованием директив открытого стандарта OpenMP [5] на системах с общей памятью.



Рис.2 – Блок-схема «Шаг Монте-Карло». В каждом блоке указаны соответствующие затраты машинного времени на их выполнение, в % от общего времени счета этого шага.

На рис. 3 представлен график коэффициента ускорения исполнения параллельной программы  $K_n = T_1/T_n$  в зависимости от количества рабочих потоков, выполняющихся на соответствующем числе вычислительных ядер системы, где  $T_n$  – время исполнения блока на  $n$  ядрах. Моделировалось осаждение 0.136 нм Ge на подложку Si размерами 13x13x4 нм со скоростью 0.00136 нм/с при температуре 450° С.

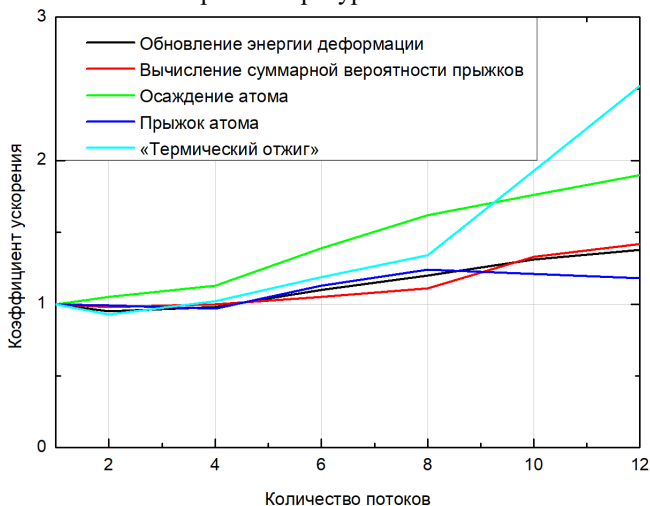


Рис.3 – Зависимость коэффициента ускорения исполнения блоков параллельной программы от количества рабочих потоков.



Наиболее эффективным оказалось распараллеливание блоков «Термический отжиг» и «Осаждение нового атома». Также можно заметить, что с ростом количества используемых потоков в подавляющем большинстве случаев коэффициент ускорения увеличивается. Тем не менее, общая эффективность распараллеливания оказалась низкой и составила от 1.25 до 3 раз в зависимости от условий модельного эксперимента. Это связано с накладными расходами, а также увеличением доли последовательных вычислений при увеличении количества осаждаемого материала.

Таким образом, были рассмотрены алгоритмы Монте-Карло моделирования гетероэпитаксиального роста. Определены вычислительно сложные блоки «Шаг Монте-Карло» и «Термический отжиг». Результаты оптимизации и распараллеливания этих блоков позволили сократить время моделирования от 1.25 до 3 раз в зависимости от условий эксперимента.

Исследование выполнено за счет гранта Российского научного фонда (проект №19-12-00070П).

#### Литература:

1. Shmidt O. G. and Eberl K. Self-assembled Ge/Si dots for faster field-effect transistors // *IEEE Trans. El. Dev.* – 2001. – Vol. 48, N 6. – P. 1175.
2. Tsybeskov L. and Lockwood D. J. Silicon-Germanium Nanostructures for Light Emitters and On-Chip Optical Interconnects // *Proc. IEEE.* – 2009. – Vol. 76, N 7. – P. 1284.
3. Novikov P. L., Nenashev A. V., Rudin S. A., Polyakov A. S., and Dvurechenskii A. V. Simulating the nucleation and growth of Ge quantum dots on Si using high-efficiency algorithms // *Nanotechnologies in Russia.* – 2015. – Vol. 10. – P. 192.
4. Rudin S. A., Zinovyev V. A., Smagina Zh. V., Novikov P. L., Nenashev A. V., Pavsky K. V. Groups of Ge nanoislands grown outside pits on pit-patterned Si substrates // *J. of Crystal Growth.* – 2022. – Vol. 593. – P. 126763.
5. Открытый стандарт для распараллеливания программ на языках Си, Си++ и Фортран. URL: <https://www.openmp.org/> (дата обращения 19.11.2023).

# ВЛИЯНИЕ ГЕОЛОГИЧЕСКИХ ПОМЕХ НА ИЗМЕРЯЕМЫЕ СИГНАЛЫ МОРСКОЙ ЭЛЕКТРОРАЗВЕДКИ С КОНТРОЛИРУЕМЫМ ИСТОЧНИКОМ В ЧАСТОТНОЙ ОБЛАСТИ

А.П. Сивенкова

Новосибирский государственный технический университет,  
г. Новосибирск, [nastya\\_sivenkova19@mail.ru](mailto:nastya_sivenkova19@mail.ru)

*В работе рассмотрено влияние геологических объектов-помех на измеряемые сигналы морской электроразведки с контролируемым источником в частотной области.*

*The article examines the influence of geological interference on the measured signals of marine controlled source electromagnetic technologies in the frequency (FD CSEM) domain.*

В морской электроразведке часто используется технология с контролируемым источником тока в частотной области (FD CSEM). Данная технология применяется для поиска углеводородов глубоко залегающих и твердых полезных ископаемых (сульфидов) на морском дне [1–2].

В работе рассматривается система Sea Bed Logging (SBL). Система SBL состоит из буксируемой генераторной линии и набора приемников, расположенных на морском дне [2].

Для моделирования электромагнитных полей используется двухэтапное выделение поля [3]. Полное трехмерное электрическое поле ищется как сумма поля  $\vec{E}^1$ , для нахождения, которого используется технология выделения поля (нормальное поле рассчитывается в горизонтально-слоистой среде), и разности полей  $\vec{E}^2$  и  $\vec{E}^3$ . Поле  $\vec{E}^3$  рассчитано в трехмерной расчетной области с искривленными границами, а поле  $\vec{E}^2$  рассчитано в трехмерной расчетной области с плоско-горизонтальными границами. Сетки для расчета полей  $\vec{E}^2$  и  $\vec{E}^3$  топологически подобные. Таким образом, полное поле  $\vec{E}'$  ищется в форме

$$\vec{E}' = \vec{E}^1 + (\vec{E}^3 - \vec{E}^2). \quad (1)$$

В свою очередь поля  $\vec{E}^1$ ,  $\vec{E}^2$  и  $\vec{E}^3$  ищутся из решения векторных уравнений:

$$\frac{1}{\mu_0} \nabla \times \nabla \times \vec{E}^3 + i\omega\sigma\vec{E}^3 = -i\omega\vec{J}^3 \quad (2)$$

$$\frac{1}{\mu_0} \nabla \times \nabla \times \vec{E}^2 + i\omega\sigma^2\vec{E}^2 = -i\omega\vec{J}^2 \quad (3)$$

$$\frac{1}{\mu_0} \nabla \times \nabla \times \vec{\mathbf{E}}^1 + i\omega\sigma^2 \vec{\mathbf{E}}^1 = (\sigma^2 - \sigma^p) \vec{\mathbf{E}}^p \quad (4)$$

где  $\mu_0$  – магнитная проницаемость вакуума,  $\sigma$  – удельная электрическая проводимость трехмерной среды,  $\sigma^2$  – удельная электрическая проводимость трехмерной среда с плоско-горизонтальными границами,  $\sigma^p$  – удельная электрическая проводимость горизонтально-слоистой среды.

С помощью рассмотренного математического аппарата было исследовано влияние объектов-помех. Геоэлектрическая модель была построена на основе месторождения Тикит в Мексиканском заливе (рис. 11111). Углеводородный резервуар представляет собой пропластки толщиной 10 метров на глубине от 1400 метров до 2000 метров с сопротивлением 10 Ом.

Рассматривалось два типа объектов помех. Первый тип - это соли (глубинные непроводящие объекты). Второй тип объектов-помех – это объекты приповерхностных слоев.

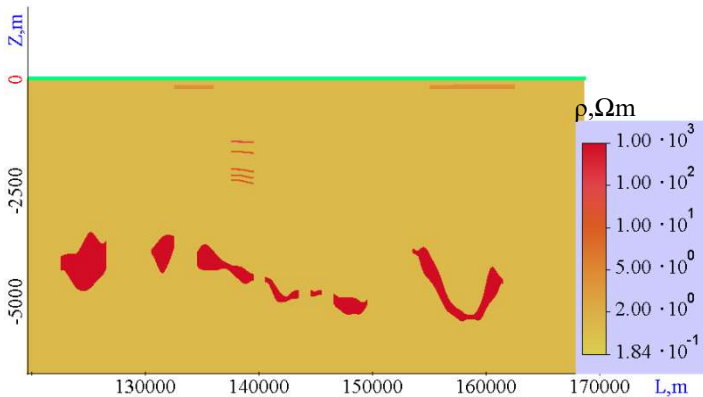


Рис. 1– Распределение сопротивления в разрезе над целевым объектом

На рис.2 представлены графики относительной разницы в сигналах на частоте 0,1 Гц (а) и на частоте 2,09 Гц (б). Из сигналов видно, что влияние от целевого объекта выше, чем влияние от соли для частоты 0,1 Гц. На частоте 2,09 Гц влияние углеводородного резервуара и соли меньше 1 процента. Влияние неоднородной части разреза на низких частотах сопоставимо с влиянием соли и целевого объекта. Но на высоких частотах отклик от верхней части разреза значительно выше, чем от солей и резервуара углеводородов. Благодаря этому, есть возможность восстановить

приповерхностные слои на высоких частотах и исключить их влияние из данных, полученных на низких частотах.

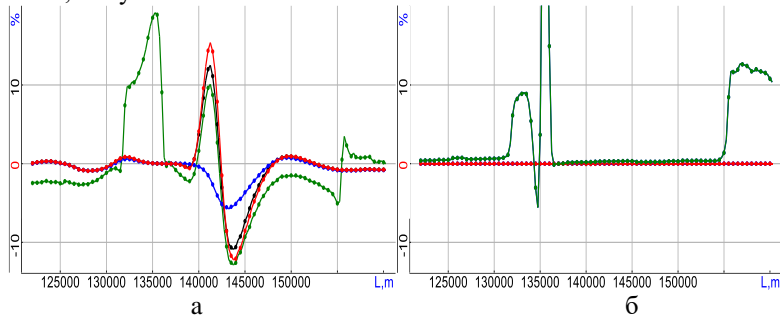


Рис. 2– Графики относительной разницы в сигналах на частоте 0,1 Гц (а) и на частоте 2,09 Гц (б) для модели, содержащей только соляные тела (синий), для модели, содержащей соль и углеводородный резервуар (черный), для модели, содержащей углеводородный резервуар (красный), и для модели, содержащей соляные тела, резервуар и неоднородности морского дна (зеленый)

Влияние объектов-помех может быть исключено из инверсии за счет использования измерений на разных частотах и специальных подходов к многоэтапной геометрической 3D инверсии.

#### Литература:

1. Marine dipole–dipole controlled source electromagnetic and coincident-loop transient electromagnetic experiments to detect seafloor massive sulphides: effects of three-dimensional bathymetry / Amir Haroon, Sebastian Hölz, Romina A S Gehrman, Eric Attias, Marion Jegen, Timothy A Minshull, Bramley J Murton // *Geophysical Journal International*. – 2018 – Vol. 215, iss. 3. – Pages 2156–2171. – <https://doi.org/10.1093/gji/ggy398>.
2. Comparison of Detection Capability by the Controlled Source Electromagnetic Method for Hydrocarbon Exploration / Guo Z, Liu J, Liao J, Xiao J. // *Energies*. – 2018. – Vol. 11, iss. 7. – 1839. – <https://doi.org/10.3390/en11071839>.
3. Geometric 2.5D inversion of marine time domain electromagnetic data with application to hydrocarbon deposits prospecting / M. G. Persova, Y. G. Soloveichik, J. K. Kjerstad, A. P. Sivenkova, A. S. Kiseleva, and D. S. Kiselev // *Journal of Applied Geophysics*. – 2023. – Vol. 212. – 104996. – <https://doi.org/10.1016/j.jappgeo.2023.104996>.

# РАЗРАБОТКА ВЫЧИСЛИТЕЛЬНЫХ СХЕМ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ С ПЕРЕНОСОМ ЧАСТИ ВЫЧИСЛЕНИЙ НА ВИДЕОКАРТУ

**Н.А. Скрынник**

**Новосибирский государственный технический университет,  
г. Новосибирск, [nikita.nstu@gmail.com](mailto:nikita.nstu@gmail.com)**

**Научный руководитель: Рояк М.Э., д.т.н., профессор**

*Статья посвящена разработке схем численного моделирования с переносом части вычислений на графический процессор. Для разработки использовались технологии NVIDIA CUDA, Numba, PyTorch, OpenCL, Vulkan и cuSOLVER. Также производилось сравнение скорости и точности работы данных схем с аналогичными схемами, предназначенными для центральных процессоров.*

*The article is devoted to the development of numerical simulation schemes with the transfer of part of the calculations to a GPU. Technologies used for development are NVIDIA CUDA, Numba, PyTorch, OpenCL, Vulkan and cuSOLVER. The performance and accuracy of these schemes have also been compared with similar schemes designed for CPUs.*

Численные методы моделирования физических процессов требуют больших вычислительных мощностей. В течение многих лет, исследователи опирались на закон Мура, согласно которому производительность центральных процессоров увеличивается вдвое каждые 2 года. Однако, современные процессоры практически достигли физических пределов своих возможностей. В связи с этим, в последнее время набирают популярность исследования, посвященные графическим процессорам, так как их архитектура полностью отличается от архитектуры центральных процессоров, и они еще не исчерпали свои возможности по увеличению производительности. Данная работа посвящена разработке вычислительных схем численного моделирования физических процессов с переносом части вычислений на видеокарту.

В ходе исследования было реализовано две схемы численного моделирования для графического процессора, а также произведено сравнение двух готовых библиотек для решения плотных и разреженных СЛАУ на центральном и графическом процессорах.

Первая схема вычисляет трехмерное граничноэлементное решение. Она была реализована на нескольких платформах для вычислений общего назначения на графическом процессоре. На рисунке 1(а) изображены результаты сравнения скорости программ, реализованных на чистом C++ на центральном процессоре и на платформах CUDA, Numba и PyTorch. Размер граничноэлементной сетки – 1000 элементов.

Из графика можно увидеть, что CUDA, Numba и PyTorch значительно обгоняют версию алгоритма для центрального процессора. Также производились сравнения скорости на сетке из 20000 элементов. Результаты представлены на рисунке 1(б). С увеличением размера сетки разрыв производительности только возрастает.

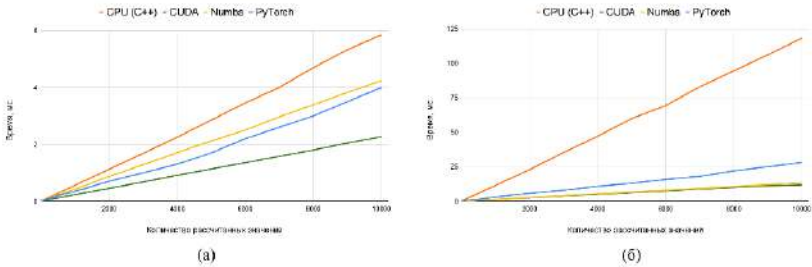


Рис. 1 – Графики времени расчета значений граничноэлементного решения в зависимости от их количества

Следующая схема, разработанная для графического процессора — это сборка граничноэлементных матриц. На рисунке 2(а) представлено сравнение скорости сборки матриц на центральном и графическом процессорах с помощью C++ и CUDA.

Также использовались технологии Vulkan и OpenCL. Сравнение их производительности с производительностью CUDA представлено на рисунке 2(б).

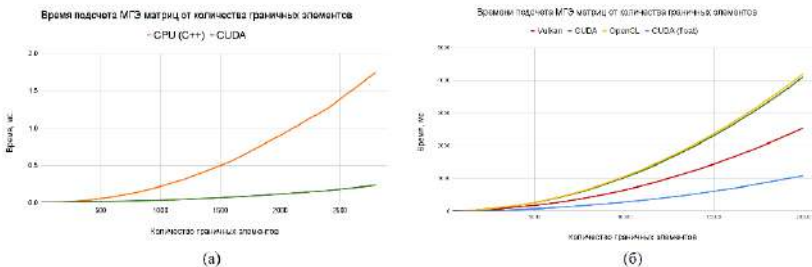


Рис. 2 – Графики времени сборки граничноэлементных матриц в зависимости от количества граничных элементов

Реализация на Vulkan работает немного быстрее CUDA и OpenCL, в связи с тем, что в данной технологии есть поддержка математических функций для чисел с плавающей запятой только одинарной точности.

Матрицы, построенные с помощью Vulkan, CUDA, использовались для решения простейшей задачи Дирихле на квадратной области

$[0, 1] \times [0, 1]$  и искомой функцией  $u(x, y) = x + y$ . На границах области использовались первые краевые условия. Размер сетки составлял 2000 элементов. Сравнение потоков функции  $p_i$  через каждый граничный элемент производилось по формуле  $1/p_i$  – точные значения потока,  $n$  – количество граничных элементов):

$$D = \frac{1}{n} \sqrt{\sum_{i=1}^n (p_i - \bar{p}_i)^2} \quad (1)$$

В таблице 1 приведены результаты сравнения точности решения задачи. Из таблицы можно увидеть, что решение задачи Дирихле с использованием матриц, собранных на Vulkan, значительно отклоняется от истинного решения в отличие от реализации на CUDA.

**Таблица 1 – Точность решения задачи Дирихле**

Платформа	Отклонение от точного решения
CUDA	$3.58285 \cdot 10^{-12}$
Vulkan	0.000106278

Также производилось сравнение библиотеки MKL для решения СЛАУ на центральных процессорах с библиотекой cuSOLVER, которая основана на платформе CUDA и позволяет решать СЛАУ на графическом процессоре. Для сравнения использовались процессор Intel Core I7 12 поколения и видеокарта RTX 3090. На рисунке 3(а) можно увидеть сравнение скорости работы данных библиотек для решения плотных СЛАУ, которые генерировались в результате применения метода граничных элементов. Библиотека cuSOLVER решает СЛАУ немного быстрее, чем MKL.

Рисунок 3(б) иллюстрирует скорость работы библиотек для решения разреженных СЛАУ, которые генерировались с помощью метода конечных элементов. Библиотека MKL работает значительно быстрее с учетом использования схемы переупорядочения строк матрицы СЛАУ перед ее факторизацией. В библиотеке cuSOLVER есть также несколько схем переупорядочения матрицы. Скорость решения СЛАУ с использованием некоторых схем представлена на рисунке. Как можно увидеть из графика ни одна из схем не позволяет решать СЛАУ также быстро как библиотека MKL. Библиотека cuSOLVER может сравниться с MKL по скорости, только если в нее передать СЛАУ с заранее переупорядоченными строками матрицы (pre-ordering).

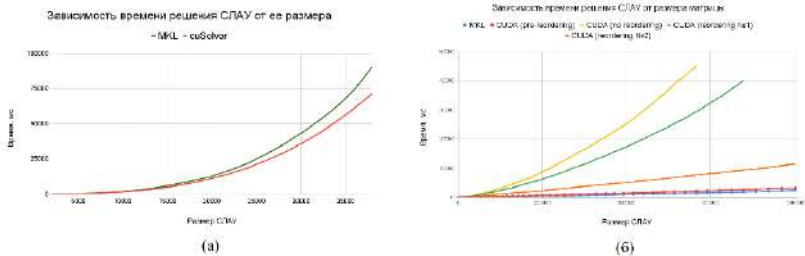


Рис. 3 – Зависимости времени решения плотных (а) и разреженных (б) СЛАУ в зависимости от ее размерности

В ходе исследования были реализованы схемы численного моделирования для подсчета граничноэлементного решения и сборки граничноэлементных матриц на графическом процессоре. Данные схемы обладают высоким потенциалом к распараллеливанию и показали значительный рост производительности по сравнению со схемами, предназначенными для центральных процессоров, в независимости от используемых технологий.

Также было произведено сравнение скорости работы библиотек cuSOLVER и MKL при решении плотных и разреженных СЛАУ. Библиотека cuSOLVER не показала значительного увеличения производительности по сравнению с MKL.

#### Литература:

1. Aamodt M., Wai Lun Fung W., Rogers T., General-Purpose Graphics Processor Architecture. – 1-е изд. – New York: Morgan & Claypool, 2018. – 122 с.
2. Hennessy J., Patterson D. Computer Architecture: A Quantitative Approach. – 5-е изд. – Waltham: Elsevier, Inc, 2012. – 1357 с.
3. Cheng J., Grossman M., McKercher T. Professional CUDA C programming. – 1-е изд. – Indianapolis: John Wiley & Sons, Inc, 2014. – 497 с.
4. Steinbach O. Numerical Approximation Methods for Elliptic Boundary Value Problems. – 1-е изд. – Wiesbaden: Springer, 2008. - 386 с.
5. Numba documentation // Numba URL: <https://numba.pydata.org/numba-doc/latest/index.html> (дата обращения: 07.02.2022).
6. Stevens E., Antiga L. Deep Learning with PyTorch: Essential Excerpts. – 1-е изд. – New York: Manning Publications Co., 2019. – 133 с.
7. General Matrix Multiple Right-Hand Side Solve // IBM: Engineering and Scientific Subroutine Library URL: [https://www.ibm.com/docs/en/essl/6.2?topic=dlaes-sgetrs-dgetrs-cgetrs-zgetrs-general-matrix-multiple-right-hand-side-solve#am5gr\\_hsgsetrs](https://www.ibm.com/docs/en/essl/6.2?topic=dlaes-sgetrs-dgetrs-cgetrs-zgetrs-general-matrix-multiple-right-hand-side-solve#am5gr_hsgsetrs) (дата обращения: 25.02.2023)



8. LAPACK Linear Equation Computational Routines // Developer Reference for Intel oneAPI Math Kernel Library for C URL: [https:// www.intel.com/content/www/us/en/docs/onemkl/developer-reference-c/2023-0/lapack-linear-equation-computational-routines.html](https://www.intel.com/content/www/us/en/docs/onemkl/developer-reference-c/2023-0/lapack-linear-equation-computational-routines.html) (дата обращения: 25.02.2023).
9. cuSOLVER API Reference // CUDA URL: <https://docs.nvidia.com/cuda/cusolver/index.html> (дата обращения: 02.03.2023).
10. Sellers G. Vulkan Programming Guide. – 1-е изд. – Boston: Addison-Wesley, 2017. – 792 с.
11. Compute Shader // Vulkan Tutorial URL: [https://vulkan-tutorial.com/Compute\\_Shader](https://vulkan-tutorial.com/Compute_Shader) (дата обращения: 08.04.2023).
12. Vulkan Memory Allocator // Github URL: <https://github.com/GPUOpen-Libraries-AndSDKs/VulkanMemoryAllocator> (дата обращения: 12.04.2023).
13. Munshi A., Gaster B., Mattson T., Fung J., Ginsburg D. OpenCL Programming Guide. – 1-е изд. – Boston: Addison-Wesley, 2012. – 603 с.

## **КОНЕЧНОЭЛЕМЕНТНЫЙ АНАЛИЗ ДЕФОРМИРОВАННОГО СОСТОЯНИЯ УПРУГОЙ ПАНЕЛИ НА РЕГУЛЯРНОЙ ПРЯМОУГОЛЬНОЙ СЕТКЕ С БИЛИНЕЙНЫМИ БАЗИСНЫМИ ФУНКЦИЯМИ**

**В.С. Слобода, М.Г. Персова**

**Новосибирский государственный технический университет,**

**г. Новосибирск, [sloboda1505@gmail.com](mailto:sloboda1505@gmail.com)**

**Научный руководитель: Персова М. Г., д. т. н., профессор**

*В статье проводится анализ результатов решения задачи о плосконапряженном состоянии упругой панели, полученных с применением метода конечных элементов (МКЭ). В ходе работы проводится создание и разработка конечно-элементной аппроксимации модели, описывающей ее плосконапряженное состояние. Завершающим этапом будет валидация созданной конечно-элементной модели (КЭМ) с вычислительным экспериментом в современных конечно-элементных пакетах.*

*The paper analyzes the results of solving the problem of plane stress state of the panel obtained using finite element method (FEM). During the work, the creation and development of finite element approximation of model describing its plane stress state is considered. Final step will be validation of the created finite element model with computational experiment in modern finite element packages.*

### Формулировка проблемы

Исследуемая панель (рис. 1) защемлена на правом торце и нагружена поверхностной силой на левом торце. Требуется создать и реализовать конечно-элементную модель, допускающую изменение всех физических и геометрических параметров.

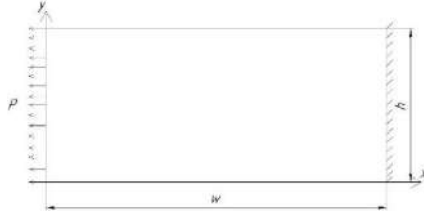


Рисунок 1 – Общий вид плоской панели

### Описание математической модели

Для двумерной задачи напряжения и деформации можно представить в виде векторов  $\sigma = \{\sigma_x, \sigma_y, \tau_{xy}\}$  и  $\varepsilon = \{\varepsilon_x, \varepsilon_y, \gamma_{xy}\}$  соответственно. В соответствии с обобщенным законом Гука для упругого изотропного тела, напряжения связаны с деформациями соотношением [1]

$$\sigma = D \cdot \varepsilon, \quad (1)$$

где  $D$  – матрица упругих постоянных, в двумерном случае имеющая вид:

$$D = \frac{E}{1-\mu^2} \begin{pmatrix} 1 & \mu & 0 \\ \mu & 1 & 0 \\ 0 & 0 & \frac{1-\mu}{2} \end{pmatrix} \quad (2)$$

Дифференциальные уравнения равновесия для двумерной задачи имеют вид:

$$\begin{cases} -\text{div}(\sigma_x, \tau_{xy}) = P_x \\ -\text{div}(\tau_{xy}, \sigma_y) = P_y, \end{cases} \quad (3)$$

где  $P_x, P_y$  – компоненты вектора поверхностных сил.

### Конечно-элементная аппроксимация математической модели

Для построения конечно-элементной аппроксимации системы (3) выбираются билинейные базисные функции  $\varphi$ , аппроксимирующие перемещения узлов элемента [2, стр. 230]:

$$\begin{aligned} \varphi_1(x, y) &= X_1(x)Y_1(y), & \varphi_2(x, y) &= X_2(x)Y_1(y) \\ \varphi_3(x, y) &= X_1(x)Y_2(y), & \varphi_4(x, y) &= X_2(x)Y_2(y), \end{aligned} \quad (4)$$

где  $X_1(x) = \frac{x_2 - x}{h_x}$ ,  $X_2(x) = \frac{x - x_1}{h_x}$ ,  $Y_1(y) = \frac{y_2 - y}{h_y}$ ,  $Y_2(y) = \frac{y - y_1}{h_y}$ ,

$$h_x = x_2 - x_1, \quad h_y = y_2 - y_1.$$

Для  $i$ -го узла конечного элемента в соответствии с методом Бубнова-Галеркина [2, стр. 63] перемещения определяются соотношением

$$\left[ \int_{\Omega} \left[ D_{11} \frac{\partial u_x}{\partial x} \frac{\partial \varphi_i}{\partial x} + D_{12} \frac{\partial u_y}{\partial y} \frac{\partial \varphi_i}{\partial x} + D_{33} \frac{\partial u_x}{\partial x} \frac{\partial \varphi_i}{\partial y} + D_{33} \frac{\partial u_y}{\partial y} \frac{\partial \varphi_i}{\partial y} \right] d\Omega = \int_s P_x \cdot \varphi_i dS \right. \quad (5)$$

$$\left. \int_{\Omega} \left[ D_{33} \frac{\partial u_x}{\partial x} \frac{\partial \varphi_i}{\partial x} + D_{33} \frac{\partial u_y}{\partial y} \frac{\partial \varphi_i}{\partial x} + D_{22} \frac{\partial u_x}{\partial x} \frac{\partial \varphi_i}{\partial x} + D_{33} \frac{\partial u_y}{\partial y} \frac{\partial \varphi_i}{\partial y} \right] d\Omega = \int_s P_y \cdot \varphi_i dS, \right.$$

где  $u_x, u_y$  - компоненты вектора  $\beta_i$  полных перемещений узла.

Вид конечного элемента и локальная нумерация его узлов представлена на рисунке 2.

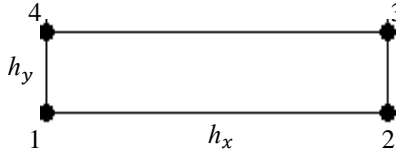


Рисунок 2 – Прямоугольный конечный элемент с локальной нумерацией узлов

При полной аппроксимации перемещений внутри элемента, т. е. поочередной замене базисных функций на соответствующие своему узлу, формируется система линейных алгебраических уравнений (СЛАУ) [3, стр. 32]

$$K \beta = R \quad (6)$$

где  $K$  – матрица жесткости всей панели,  $\beta$  - вектор перемещений узлов,  $R$  - вектор нагрузок.

Таким образом, решением системы (6) будет являться вектор перемещений, определяющий деформированное состояние панели.

#### Валидация конечно-элементной модели

В качестве подтверждения правильности созданной конечно-элементной модели предлагается сравнить перемещения нескольких узлов нагруженного торца, полученные из решения СЛАУ (6) с перемещениями, полученными из решения задачи в программных комплексах FEMAP и ANSYS, на одинаковых сетках [4, 5].

Для валидации реализованной КЭМ проводится серия расчетов пластины с различными размерами, сеткой и нагрузкой. Результаты валидации приведены в таблице 1.

Деформированное состояние панели из первого расчета представлено на рисунке 4.

**Таблица 1 – Валидация КЭМ с моделями в комплексах**

№ рас- чета	h, мм	w, мм	Толщи- на t, мм	Число разби- ний		Перемещения первого узла, мм			Разница %
				по x	по y	КЭМ	FEMAP	ANSYS	
Материал – Ст3, нагрузка P <sub>x</sub> = 11000 Н/м									
1	30	50	3	10	10	0,02885	0,02886	0,02886	0,035
2	60	100	1	10	10	0,08656	0,08658	0,08658	0,023
Материал – Ст3, Нагрузка P <sub>x</sub> = 33000 Н/м									
3	60	100	2	25	15	0,12990	0,13000	0,13000	0,770
4	80	50	2	25	15	0,04994	0,04993	0,04993	0,020
Материал – Д16Т, нагрузка P <sub>x</sub> = 50000 Н/м									
5	30	50	3	20	20	0,38218	0,38230	0,38230	0,031

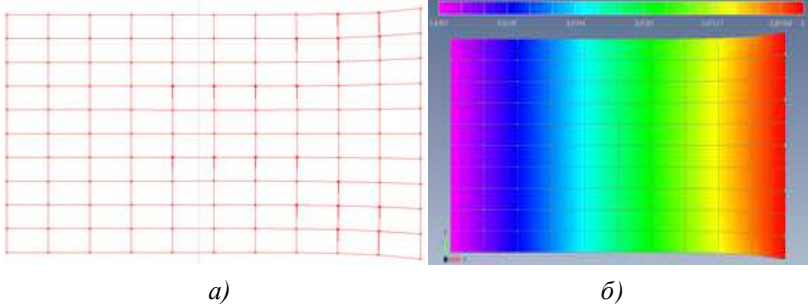


Рисунок 4 – Деформированное состояние панели, полученное а) – реализацией КЭМ на C++, б) – расчетом в среде FEMAP

### Итоги анализа

Приведенная в таблице 1 разница между результатами написанной программы и пакетами ANSYS и FEMAP подтверждает пригодность первой для расчета и анализа плосконапряженного состояния панели с различными характеристиками материала, параметрами геометрии и конечно-элементной сетки. Также имеется перспектива доработки программы до наличия возможности обработки геометрии с помощью кривых и моделирования анизотропии материала.

### Литература:

1. Демидов С. П. Теория упругости: Учебник для вузов. – М.: Высш. школа, 1979. – 432 с., ил.
2. Ю. Г. Соловейчик, М. Э. Рояк, М. Г. Персова. Метод конечных элементов для решения скалярных и векторных задач: учеб. пособие. Новосибирск: Изд-во НГТУ, 2007. – 896 с. («Учебники НГТУ»).
3. Присекин В. Л., Расторгуев Г. И. Основы метода конечных элементов: Учеб. пособие. – Новосибирск: Изд-во НГТУ. -2003. – 124 с.

4. Рычков С. П. Моделирование конструкций в среде Femap with NX Nastran. – М.: ДМК Пресс, 2013. – 784 с.: ил.

5. А. В. Чигарев, А. С. Кравчук, А. Ф. Смалюк. ANSYS для инженеров: Справочное пособие // Машиностроение-1. — М. – 2004. 512 с.

## **ИССЛЕДОВАНИЕ ВЛИЯНИЯ РЕГУЛЯРИЗАЦИИ НА РЕШЕНИЕ НЕЛИНЕЙНОЙ ОБРАТНОЙ ЗАДАЧИ МАГНИТОРАЗВЕДКИ**

**Л.А. Столярова**

**Новосибирский государственный технический университет  
г. Новосибирск, lili-stol@yandex.ru**

**Научный руководитель: Соловейчик Ю.Г., д.т.н., профессор**

*В данной работе рассматривается влияние регуляризации на решение обратной задачи магниторазведки при использовании геометрической параметризации и метода Гаусса-Ньютона для трехмерных объектов.*

*The paper describes methods of regularization for solving the inverse task of magnetic survey with geometric parametrization and Gauss-Newton method for detection of 3-D objects.*

**Введение.** Магниторазведка является геофизическим методом, направленным на восстановление в исследуемой области местоположения геологических объектов, обладающих остаточной намагниченностью.

Решение обратной задачи, или инверсия практических данных, заключается в восстановлении параметров по данным прямой задачи – сигналам из приёмных установок, расположенных на поверхности земли.

Результатом решения обратной задачи является восстановленная среда, магнитное поле которой близко к его измеренным значениям в приёмниках. В результатах геофизической инверсии повсеместно встречаются неопределенность и неединственность решения. Для того чтобы избавиться от неоднозначности результатов инверсии и получить более точное решение, используется регуляризация обратной задачи [2].

**Математическая модель.** Восстановление геометрических и физических параметров среды является нелинейной обратной задачей. В данной работе используется геометрическая параметризация: модель задается количеством объектов с некоторой намагниченностью и их положением в расчетной области.

Для нахождения геометрических параметров и значений намагниченности в обратной задаче применяется метод Гаусса-Ньютона. Вектор параметров начального приближения задается формулой (1), где геометрические параметры задаются как минимальная и максимальная координата по каждому из измерений (x, y, z).

$$u = (x_{min}, x_{max}, y_{min}, y_{max}, z_{min}, z_{max}, P_x, P_y, P_z) \quad (1)$$

Значения параметров определяются на основе минимизации функционала, представляющего собой сумму взвешенных квадратов невязки практических ( $B_i$ ) и теоретических ( $\tilde{B}_i$ ) данных:

$$\Phi = \sum_{i=1}^N \omega_i^2 [(B_i^x - \tilde{B}_i^x)^2 + (B_i^y - \tilde{B}_i^y)^2 + (B_i^z - \tilde{B}_i^z)^2] \quad (2)$$

где  $\omega_i$  – веса, отражающие уровень погрешности,  $N$  – количество приёмников [1].

Для решения задачи используется следующая регуляризация: ограничение резкого изменения значений параметров (как геометрических, так и физических), а также предотвращение пересечения границ.

**Исследования.** Отсутствие ограничения на пересечение границ по каждой из координат делает возможным возникновение ситуации, при которой минимальная и максимальная границы меняются местами.

На рисунке (1) приведено решение обратной задачи с применением регуляризации. Выполнялась попарная проверка положения границ по одной координате относительно друг друга. Значения намагниченности задавались равными истинным и не изменялись в процессе подбора параметров. Истинная аномалия показана зеленым цветом, начальное приближение – синим контуром, восстановленная модель - красным контуром. Решение сходится к истинному положению аномального объекта.

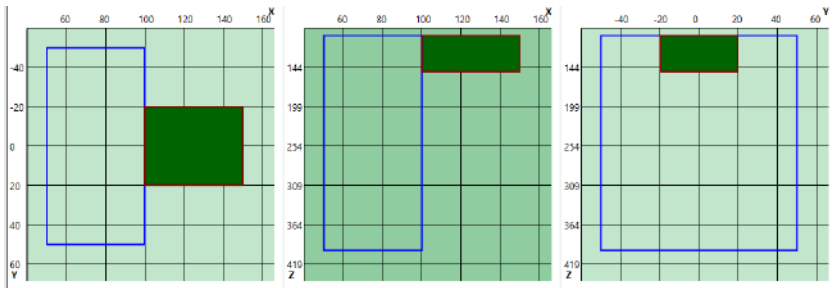


Рис. 1 – Решение обратной задачи с применением регуляризации геометрических параметров

При решении этой же задачи без регуляризации уже на 4-ой итерации значение левой границы  $y_{\min}$  становится больше, чем значение правой границы  $y_{\max}$ , после чего решение начинает расходиться, параметры показаны в таблице (1).

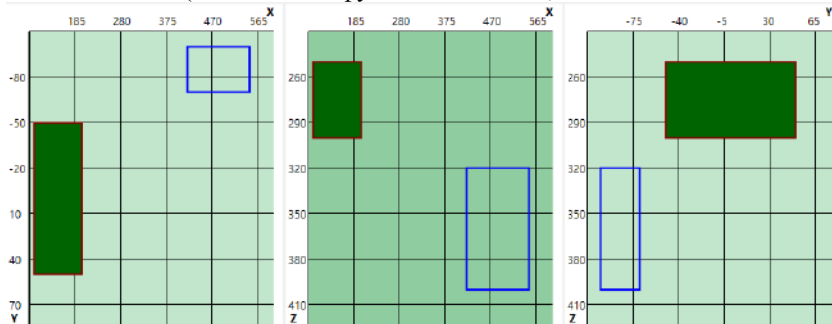
**Таблица 1 – Решение обратной задачи без регуляризации геометрических параметров**

	3 итерация	4 итерация	7 итерация
$x_{\min}$	88.3634	101.477	<b>67.14</b>
$x_{\max}$	103.205	142.64	<b>29.97</b>
$y_{\min}$	-10.9753	<b>5.4933</b>	<b>446.88</b>
$y_{\max}$	57.9906	<b>-3.53271</b>	<b>-164.72</b>
$z_{\min}$	-213.242	-114.605	<b>445.87</b>
$z_{\max}$	-101.4	-107.829	<b>-261.69</b>

В случае одновременного поиска и координат, и значений намагниченности отсутствие ограничений на максимальное изменение параметров намагниченности, может привести к их резкому возрастанию/понижению. Если объем истинной аномальной области и начального приближения сильно различаются, разница объёмов может начать компенсироваться изменением значения намагниченности вместо подбора координат.

Для того чтобы избежать этого, выполняется следующая проверка: значение параметра намагниченности между итерациями не должно изменяться более чем на допустимый процент относительно значения этого параметра на предыдущей итерации.

На рисунке (2) приведено решение обратной задачи с применением регуляризации – запрет изменения намагниченности более чем на 50%. Найденное решение сходится к истинному положению аномалии с заданной точностью (взвешенный функционал  $< 10^{-3}$ ).



**Рис. 2 – Решение обратной задачи с применением регуляризации физических параметров**

В таблице (2) представлено соответствие количества итераций, за которое было найдено решение, и заданных допустимых изменений параметра между итерациями.

**Таблица 2 – Влияние ограничений на количество итераций инверсии**

Допустимое изменение параметра относительно значения на предыдущей итерации, %	20	50	100	200
Количество итераций	34	26	32	40

Отсутствие регуляризации приводит к резкому изменению значений намагниченности (Таблица 3).

**Таблица 3 – Решение обратной задачи без регуляризации физических параметров**

	Начальное значение	1 итерация	2 итерация
$p_x$	0.2	6.12318	<b>-95.4778</b>
$p_y$	0.2	6.25928	<b>-89.5573</b>
$p_z$	0.2	6.29341	<b>-81.0043</b>

Ещё один метод корректировки задачи заключается в добавлении некоторой малой величины на диагональ матрицы при решении СЛАУ. Это необходимо для того, чтобы избежать вырожденности СЛАУ, так как в ходе итерационного процесса подбора параметров значения намагниченности могут принимать значения, близкие к нулю.

**Заключение.** Применение регуляризации в методе Гаусса-Ньютона позволяет сделать поиск параметров более устойчивым и более точно выделить положение аномалий в расчётной области.

Отсутствие регуляризации приводит к невозможности нахождения истинного положения аномалии даже для простых случаев: один объект в расчетной области, подбор только геометрических параметров с фиксированной истинной намагниченностью, задание начального контура близко к истинному положению аномалии.

Литература:

1. Вагин Д.В. Оценивание параметров в обратных задачах: учебное пособие / Д.В. Вагин. – Новосибирск: Изд-во НГТУ, 2019. – 48 с.
2. Персова М.Г. Современные компьютерные технологии: конспект лекций / М.Г. Персова, Ю.Г. Соловейчик, П.А. Домников. – Новосибирск: Изд-во НГТУ, 2014. – 80 с.



# РАЗРАБОТКА МОДЕЛИ НЕЛИНЕЙНОГО МУЛЬТИПОЛЬНОГО КИКЕРА ДЛЯ ДОБАВОЧНОЙ ИНЖЕКЦИИ В НАКОПИТЕЛЬНОЕ КОЛЬЦО КОМПЛЕКСА СКИФ

**И.А.Ульев, В.Д. Кашкин**

**Институт ядерной физики им Г. И. Будкера СО РАН,  
г. Новосибирск**

**Новосибирский государственный технический университет,  
г. Новосибирск, i.ulev1998@mail.ru**

**Научный руководитель: Бурдаков А.В., д. ф.-м. наук, профессор,  
Левичев Е.Б., д. ф.-м. наук, профессор**

*Особенностью проекта СКИФ является зарекомендованный малый эмиттанс пучка порядка  $75 \text{ пм}^* \text{рад}$ . В источниках синхротронного излучения (СИ) последних поколений для достижения малого эмиттанта пучка используются ТМЕ-ячейки (Theoretical Minimum Eittance) различных конфигураций. Использование сильных секstupольных линз для подавления хроматизма значительно уменьшает площадь динамической апертуры. Это может существенно затруднять инжекцию в накопительные кольца. В связи с этим, разработка идей связанных с модернизацией методов инжекции является актуальной научной задачей. В качестве методов модернизации системы инжекции авторы доклада предлагают использовать нелинейный мультипольный кикер (НМК), размещённый в магнитной структуре специальным образом. В настоящей работе в среде программного комплекса MAD-X исследована инжекция в накопительные кольца с малой динамической апертурой при помощи НМК. Кроме того, представлена модель и конструкция НМК разработанная по примеру модели кикера системы добавочной инжекции разработанных ведущими лабораториями Японии (КЕК-PF) и Швеции (MAX IV).*

*A feature of the SKIF project is the proven small beam emittance of about  $75 \text{ pm}^* \text{rad}$ . In synchrotron radiation (SR) sources of recent generations, TME cells (Theoretical Minimum Emission) of various configurations are used to achieve a small beam emittance. The use of strong sextupole lenses to suppress chromaticism significantly reduces the area of the dynamic aperture. This can significantly complicate injection into the storage rings. In this regard, the development of ideas related to the modernization of injection methods is an urgent scientific task. As methods of modernization of the injection system, the authors of the report suggest using a nonlinear multipole kicker (NMK) placed in a magnetic structure in a special way. In this paper, injection into storage rings with a small dynamic aperture using NMK is investigated in the environment of the MAD-X software package. In addition, the model and design of the NMK developed by the example of the kicker model of the additive injection system developed by leading laboratories in Japan (KEK-PF) and Sweden (MAX IV) are presented.*

Идея НМК состоит в том, что удар кикера уменьшает угловую координату. С этой точки зрения нас интересует поиск места  $\pi/2$  по фазе (где

координата минимальна). Задача о расположении НМК сводится к поиску оптимального места с точки зрения набег фазы и оптических функций Твисса, однако стоит учитывать возможности технической реализации создания необходимого градиента поля кикера, а также требования не возбудить колебания накопленного пучка так как это негативно скажется на пользователях станций СИ.

### Моделирование в среде MAD-X [1]

Так как НМК является однооборотным устройством имеет смысл проводить моделирование первого оборота после инжекции, для определения его действия на динамику в течении первого оборота.

Моделирование будем проводить по следующему алгоритму:

1. Расположив НМК в предполагаемое место, определим необходимое поле, пройдя диапазон с необходимым шагом, определим пик его эффективности.

2. Методом Симпсона найдем интеграл модуля траектории пучка без удара кикера и с ударом кикера. Отношение площадей будем считать за эффективность удара НМК.

3. Изобразим траектории для сравнения движения пучка, а также изобразим движение пучка в нормализованном фазовом пространстве до удара и после удара.

Первое место выберем в области 19.6 (разумеется, от точки инжекции) так как, там имеется достаточно большое отклонение пучка от орбиты (4 мм). Более того, это начало прямолинейного промежутка, а значит, можно удобно разместить устройство.

Представлены график зависимости эффективности от поля кикера, а также удар НМК в нормализованном фазовом пространстве.

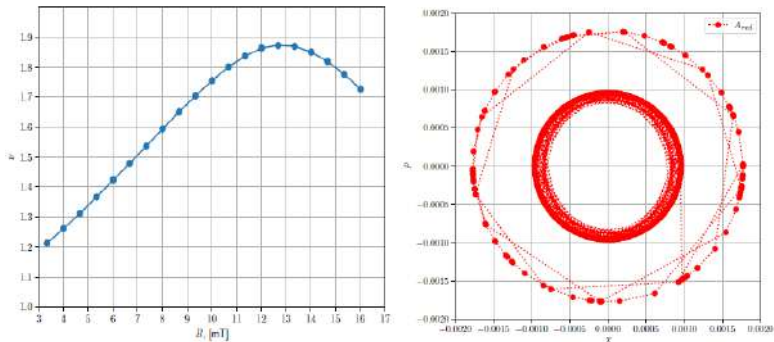


Рис.1 – Слева показана зависимость эффективности от величины поля, справа - удар в нормализованной фазовой плоскости.

Также представлены траектория пучка до и после удара кикера. Для такого удара необходимо создать поле в 12.6 мТ, при отклонении пучка в 4 мм.

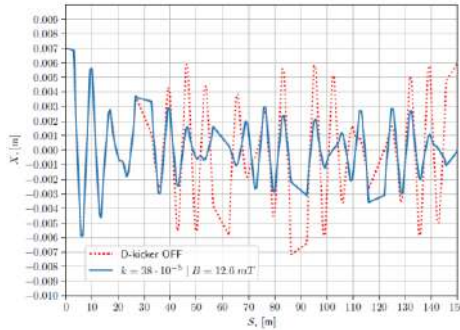


Рис.2 – Красным цветом обозначена траектория до удара. Синим - после удара НМК.

### Разработка конструкции НМК

На этапе разработки нелинейного мультипольного кикера было принято решение отталкиваться от успешного опыта ведущих лабораториями Японии (КЕК-PF) и Швеции (MAX IV) [2, 3].

В ходе инжекции, на инжектируемый пучок должно воздействовать магнитное поле, необходимое для уменьшения амплитуды колебаний инжектируемого пучка относительно накопленного и при этом оно не должно оказывать существенного влияния на накопленный пучок. Такая форма поля достигается за счет подбора различных комбинацией квадрупольной, секступольной и октупольной составляющих магнитного поля.

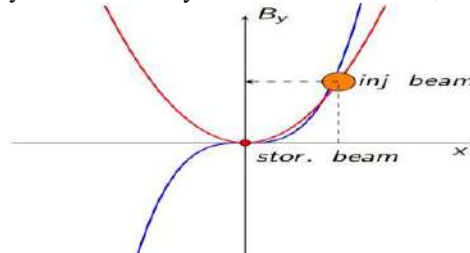


Рис.3 – Схема добавочной инжекции с использованием НМК в накопительное кольцо.

Требуемое магнитное поле при отклонении инжектируемого пучка на  $\alpha=0.380$  мрад мм обеспечивается следующей длиной магнита:

$$L = \frac{\alpha \cdot B \rho}{B} = 0.3 \text{ м}$$

Где  $B\rho$  - магнитная жесткость системы  $10 \text{ Тл}^*\text{м}$ ,  $B$  – магнитное поле равное  $12.6 \text{ мТ}$ .

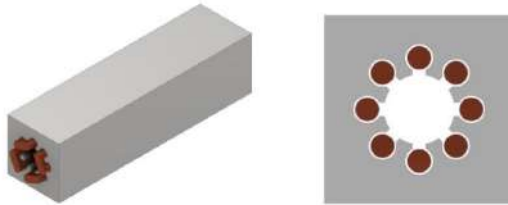


Рис.4 - Слева представлен общий вид НМК, справа – поперечное сечение

Результат моделирования в среде Comsol Multiphysics:

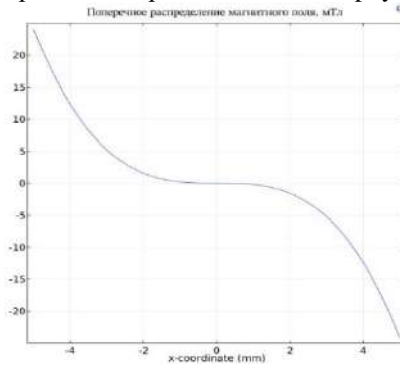


Рис.5 - Поперечное распределение магнитного поля.

### Заключение

Исходя из полученного распределения видно, что в области накопленного пучка магнитное поле принимает значение близкое к нулю, что вызовет незначительные возмущения накопленного пучка. Также видно, что в точках  $x=\pm 4\text{мм}$ , амплитуда соответствует  $12.5 \text{ мТл}$ , что близко к рассчитанному ранее значению поля.

На основе созданной компьютерной модели и исходя из полученных результатов, возможна реализация реального устройства НМК.

### Литература

1. MADX, <http://madx.web.cern.ch/madx>, Tech. Rep.
2. Kentaro Harada, Yukinori Kobayashi, Tsukasa Miyajima and Shinya Nagahashi PF-AR INJECTION SYSTEM WITH PULSED QUADRUPOLE MAGNET // Proceedings of APAC. - 2004. - С.344-346.
3. S.C.Leemann Pulsed sextupole injection for Sweden's new light source MAX IV // PHYSICAL REVIEW SPECIAL TOPICS – ACCELERATORS AND BEAMS. - 2012. - №15. - С. 050705-1 - 050705-20.

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО  
ИНТЕЛЛЕКТА, МАШИННОГО ОБУЧЕНИЯ И  
ПРЕДСКАЗАТЕЛЬНОГО МОДЕЛИРОВАНИЯ**

**ОБ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ РАБОТЫ ГЛУБОКИХ  
НЕЙРОННЫХ СЕТЕЙ ДЛЯ ЯЗЫКОВЫХ МОДЕЛЕЙ**

**А.Д. Карасюк**

**Новосибирский государственный технический университет,  
г. Новосибирск, karasyuk.2018@stud.nstu.ru**

**Научный руководитель: Авдеенко Т.В., д.т.н., профессор**

*В настоящей работе проводится анализ методов интерпретации глубоких нейронных сетей применительно к языковым моделям типа GPT. Мы предлагаем подход, основанный на внутренних представлениях, подобных человеческому мышлению. Выявлены ключевые аспекты обучения с использованием модели обучения с подкреплением RLHF. Рассматриваются стратегии контроля обучения на разнообразных данных, подчеркивается важность изучения внутренних представлений для глубокого понимания модели.*

*This work presents an analysis of deep neural network interpretation methods, specifically applied to language models like GPT. We propose an approach based on internal representations resembling human thought processes. Key aspects of training using the RLHF reinforcement learning model are identified. Learning control strategies are examined across diverse datasets, emphasizing the significance of studying internal representations for a profound understanding of the model.*

Целью настоящей работы является исследование и поиск новых методов для интерпретации глубоких нейронных сетей применительно к интерпретации больших языковых моделей для генерации текстов Generative Pre-trained Transformer (GPT). В данном случае мы отходим от классической низкоуровневой интерпретации и абстрагируемся до уровня представлений, подобно представлениям человека.

Для моделей Computer Vision можно рисовать промежуточные признаки, чтобы понять, на что именно обращает внимание модель. В нашем случае они будут интерпретированы нами как некоторые шаблоны, которыми оперирует модель. В задаче генерации новых картинок можно выделить важные образы для генерации. В случае моделей NLP можно исследовать поведение attention слоев, понимая, какие представления слой выучивает (например, отношение подлежащее-сказуемое). С помощью методов типа SHAP можно отслеживать, что больше всего повлияло на генерацию следующего слова.

Однако важно понимать, что представляет собой результат работы модели (результат генерирования текста). При исследовании слоев внимания (attention-слоев) для базовой языковой модели мы можем проследить, как она изучает язык, и выявить, какие отношения вообще существуют между словами (рис. 1). В результате мы убеждаемся, что модель выучила язык и понимает его структуру для генерации непротиворечивых грамматических конструкций.

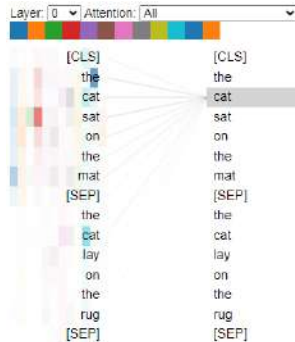


Рис. 1 – Иллюстрация изображения промежуточных слоев внимания внутри модели BERT

На Рисунке 1 показана связь слова *cat* со всеми остальными. Слева отображаются слои внимания, где прозрачностью прямоугольника обозначена сила активации для комбинации слов в левой колонке со словом *cat*. CLS – обозначения токена для классификации. SEP – обозначение для токена разделителя частей предложения.

При дообучении с помощью подхода Reinforcement Learning with Human Feedback (RLHF) модель может выучить более сложные конструкции, полученные в результате обратной связи от людей, такие как, например, специфичные стили, тональность повествования, контекстуальные особенности (упоминание тем, объектов и событий), решение специфических задач [1,2].

Из-за разнообразия текстов в наборах данных, предназначенных для обучения языковых моделей, с различными целями, модель также неявно усваивает это разнообразие. С одной стороны, для обучения желательно использовать как можно больше разнообразных данных. С другой стороны, получить хороший набор, где потенциально не будет проблем (любого характера, например, искажение фактов, большое количество рекламного спама, призывами к насилию), крайне трудно, т.к. это требует большого количества ресурсов. Из-за наличия такого рода наборов отсутствует контроль над тем, чему модель научилась. Это, в свою

очередь, дает простор для взломов моделей (при взломе модели могут генерировать текст, который может навредить или который будет социально не одобряем и т.д.), даже при сильных ограничениях (как у ChatGPT) [3]. От всех проблем избавиться невозможно, однако хотелось бы следить за некоторыми показателями в реальном времени, которые могут говорить о том, что на данный момент модель сильно отклонилась от желаемого поведения. Последнее также можно использовать для контроля генерации текста в реальном времени.

Поэтому важно исследовать такие аспекты генерируемого текста, как правдивость, полезность, вероятность появления, мораль и эмоции, а также способы выявления таких процессов, как стремление к власти или обману. Для этого мы можем прибегнуть к концепции изучения представлений модели. Используя подготовленные данные для каждого аспекта текста и процесса, мы можем найти направления (так как токены внутри модели представлены векторами большой размерности), в которых могут находиться внутренние представления токенов в скрытых слоях (например, для отслеживания наличия таких кластеров мы можем отображать скрытые представления с помощью методов снижения размерности данных PCA и других) [4,5]. Мы можем отслеживать каждое представление и, в случае обнаружения некорректного аспекта или процесса с точки зрения применимости текста, корректировать его от слоя к слою.

Исследование внутренних представлений модели предоставляет возможность глубже понять, какие семантические концепции и отношения формируются на различных уровнях обработки информации. Этот подход позволяет нашей методологии перейти от простого слежения за признаками к анализу смысловых закономерностей внутри модели. Например, мы можем дополнительно исследовать, какие типы информации влияют на активации определенных нейронов в различных слоях, что позволит более точно определить, какие аспекты контента оказывают наибольшее влияние на принятие решений моделью.

С использованием современных методов, таких как вариационные автоэнкодеры и обучение на граничных данных (adversarial training), мы также можем стремиться к более точной настройке представлений модели [6]. Это не только повысит ее эффективность в решении конкретных задач, но и сделает процесс обучения более интерпретируемым.

Важным направлением для будущих исследований является изучение взаимодействия различных типов данных в процессе обучения. Комбинирование множества задач источников данных может дать новые ключи для понимания, как модель выстраивает свои внутренние

представления в зависимости от контекста. Дополнительно можно рассмотреть внедрение механизмов обучения с подкреплением, способных корректировать внутренние представления на основе обратной связи от конечного пользователя. Это создаст возможность более гибкого и персонализированного взаимодействия с моделью.

Вместе с тем, интеграция методов, подобных тем, что представлены в техническом отчете GPT-4, может добавить новые уровни гибкости и контроля, предоставляя инструменты для настройки работы модели под конкретные потребности и этические стандарты [1]. Наконец, изучение представлений модели и их коррекция в реальном времени открывают перспективы для создания более интеллектуальных и этических систем генерации текста, поддерживая важные аспекты правдивости, полноты и эмоциональной поддержки в общении.

В итоге рассмотрения вопросов интерпретации глубоких нейронных сетей, особенно в контексте языковых моделей типа Generative Pre-trained Transformer (GPT), мы выделили ключевые аспекты прозрачности и контроля. Уходя от традиционных методов интерпретации, мы применили подход, аналогичный внутренним представлениям человека.

Используя методы визуализации промежуточных признаков и анализа слоев внимания, мы стремились не только понять, как модель работает, но и обеспечить контроль и направление ее деятельности. Стоит отметить, что данное исследование не только предоставляет обзор текущего состояния проблемы, но и поднимает важные вопросы доверия, предсказуемости и этической приемлемости сгенерированных текстов. Эти аспекты особенно актуальны в свете быстрого развития и распространения языковых моделей.

Нам предстоит дальнейшее углубление в концепцию изучения представлений модели, применение методов снижения размерности, таких как PCA, а также использование обратной связи от пользователей. Все это поможет не только повысить интерпретируемость, но и разработать более эффективные методы контроля за генерацией текста.

Таким образом, данное исследование поможет лучше контролировать результат генерирования текста. Помимо того, что мы имеем возможность “направлять” модель, мы также сможем программировать некоторый алгоритм, по которому она будет взаимодействовать с конечным пользователем. При отклонении от алгоритма модель прервет диалог. В данном случае работа модели будет намного прозрачнее, что повысит ее предсказуемость и, как следствие, интерпретируемость.



Литература:

1. *OpenAI GPT-4 Technical Report* // arxiv – 2023, March. – URL: <https://arxiv.org/pdf/2303.08774.pdf>
2. *Paul Christiano, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, Dario Amodei*. Deep reinforcement learning from human preferences // arxiv – 2017, June. – URL: <https://arxiv.org/pdf/1706.03741.pdf>
3. *Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright* Training language models to follow instructions with human feedback // *NeurIPS Proceedings*, – 2022, November – URL: [https://proceedings.neurips.cc/paper\\_files/paper/2022/hash/b1efde53be364a73914f58805a001731-Abstract-Conference.html](https://proceedings.neurips.cc/paper_files/paper/2022/hash/b1efde53be364a73914f58805a001731-Abstract-Conference.html)
4. *Stephanie Lin, Jacob Hilton, Owain Evans* TruthfulQA: Measuring How Models Mimic Human Falsehoods // arxiv – 2021, September – URL: <https://arxiv.org/pdf/2109.07958.pdf>
5. *Andy Zou, Long Phan, Sarah Chen, James Campbell* Representation Engineering: A Top-Down Approach to AI Transparency // arxiv – 2023, October – URL: <https://arxiv.org/pdf/2310.01405.pdf>
6. *Ali Shafahi, Mahyar Najibi, Amin Ghiasi* Adversarial Training for Free! // arxiv – 2019, April – URL: <https://arxiv.org/pdf/1904.12843.pdf>

## **МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ СИГНАЛОВ ЭЛЕКТРОЭНЦЕФАЛОГРАММ ДЛЯ ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЕЙ КОГНИТИВНЫХ ПРОЦЕССОВ**

**А.А. Карманова<sup>1</sup>, Е. А. Сапина<sup>2</sup>**

**<sup>1</sup>Новосибирский государственный технический университет,  
г. Новосибирск, [karmanova.2023@stud.nstu.ru](mailto:karmanova.2023@stud.nstu.ru)**

**<sup>2</sup>Новосибирский государственный университет экономики и  
управления, г. Новосибирск, [cdoastrum@mail.ru](mailto:cdoastrum@mail.ru)**

**Научный руководитель: Карманов В. С., к.т.н., доцент**

*Для исследования нейрофизиологических механизмов функционирования нервной системы и психики, применяются современные компьютерные технологии и методы математической обработки данных. Одним из неинвазивных методов определения функционального состояния головного мозга является электроэнцефалография. Для определения показателей когнитивных процессов предлагается использование метода спектрального сингулярного анализа, совместно с традиционным подходом, использующим Фурье-анализ.*

*To study the neurophysiological mechanisms of the functioning of the nervous system and psyche, modern computer technologies and methods of mathematical data processing are used. One of the non-invasive methods for determining the functional*

*state of the brain is electroencephalography. To determine indicators of cognitive processes, it is proposed to use the method of spectral singular analysis, together with the traditional approach using Fourier analysis.*

Когнитивные расстройства и фоновые состояния личности тесным образом связаны с психофизиологическими показателями. Еще в 1929 Хансом Бергером была впервые зарегистрирована электроэнцефалографическая активность головного мозга примерно на частоте в 10 Гц. Открытый ритм получил название альфа-ритма [1]. В дальнейшем было показано, что именно альфа-ритм связан со многими процессами центральной нервной системы: состоянием фоновой активности, с произвольной и непроизвольной регуляцией поведения, решением когнитивных задач, процессами внимания [2-4].

Согласно результатам многочисленных исследований альфа-ритма электроэнцефалограмм (ЭЭГ), выделяют три основные характеристики альфа-активности: частота доминирующего ритма ЭЭГ, реакция снижения амплитуды на открывание глаз и веретенообразность колебаний [5-7].

Характеристики альфа-ритма являются индикаторами общего состояния человека. Так, например, частота альфа-активности связана со скоростью воспроизведения вспоминаемой информации [8-10], наилучшая переключаемость наблюдается у людей с широким альфа-диапазоном [11], более длительные веретенообразные участки альфа-ритма свидетельствуют о большей эмоциональной устойчивости и уравновешенности индивидов [12].

В настоящее время задача по разработке и исследованию математических методов анализа индивидуального сочетания нескольких характеристик альфа-активности ЭЭГ, имеющего определенную физиологическую значимость, продолжает оставаться востребованной и актуальной [13, 14]. Разрабатываются новые подходы, модели и методы, позволяющие уточнять или получать больше информации из сигналов ЭЭГ [15, 16].

Традиционно для анализа ЭЭГ применяется дискретное преобразование Фурье (ДПФ). Недостатком ДПФ является условие стационарности исходного сигнала, что в случае ЭЭГ почти никогда не выполняется. В связи с этим применяются различные варианты и модификации этого метода, связанные, например, с применением оконного ДПФ, усреднением, сглаживанием спектра на различных интервалах, фильтрацией шума и т.п. [14]. На рис. 1 приведен фрагмент измеренного сигнала ЭЭГ (первая секунда) для случая открытых и закрытых глаз испытуемого, на рис. 2 – фрагмент амплитудного спектра ДПФ полученных сигналов в полосе от 0 до 50 Гц.

В данной работе предлагается в сочетании с традиционным подходом, использующим Фурье-анализ, также применять спектральный сингулярный анализ (ССА, SSA, Spectral Singularity Analysis). Метод ССА позволяет определять и раскладывать исходный сигнал на отдельные аддитивные составляющие. Такими составляющими могут быть трендовые компоненты, гармонические (сезонные) колебания, шум и пр. В отличие от работы [17], использующей схожий подход, проводится постобработка результатов, позволяющая определять частотные характеристики компонент [18]. В силу нестационарности сигнала, на каждый отсчет сигнала вычисляется оценка мгновенной частоты, и далее делается усреднение (или выбор медианного значения), которое и принимается за оценку частоты каждой отдельной компоненты.

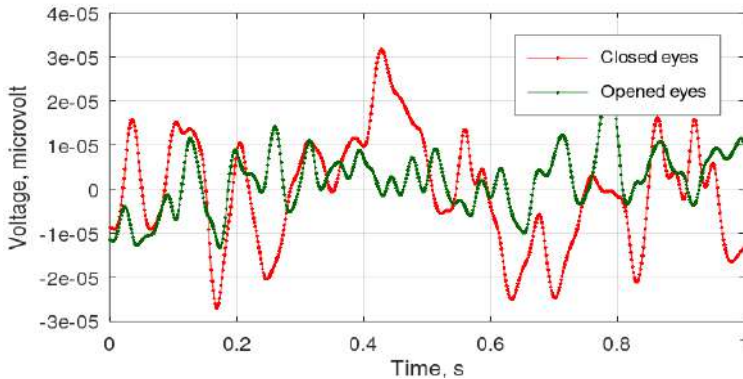


Рис. 1 – Фрагмент ЭЭГ (первая секунда), при открытых и закрытых глазах испытуемого

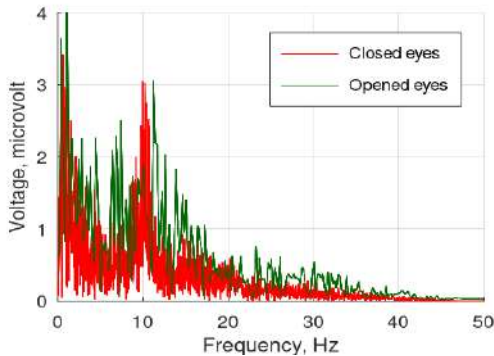


Рис. 2 – Фрагмент амплитудного спектра сигнала

Алгоритм обработки сигналов ЭЭГ, с адаптаций применения метода ССА, предлагается следующим:

Выбор настраиваемых параметров метода ССА (ширина окна, определяет траекторную размерность сигнала).

Формирование траекторной матрицы ряда/сигнала.

Сингулярное разложение траекторной матрицы, определение сингулярных значений и матриц левых и правых сингулярных векторов, разложение траекторной матрицы на элементарные.

Вычисление элементарных компонент (диагональное усреднение/ганкелизация элементарных матриц).

Анализ и формирование отдельных компонент (группировка элементарных компонент со схожими частотными характеристиками, отбрасывание низкоамплитудных, шумовых компонент и т.д.).

Анализ частотных характеристик отдельных компонент, отбраковка аномальных наблюдений, оценка частоты (по выборочному среднему или выборочной медиане).

Определение показателей когнитивных процессов (частоту максимальной гармоники, ширину диапазона для открытых/закрытых глаз и др.).

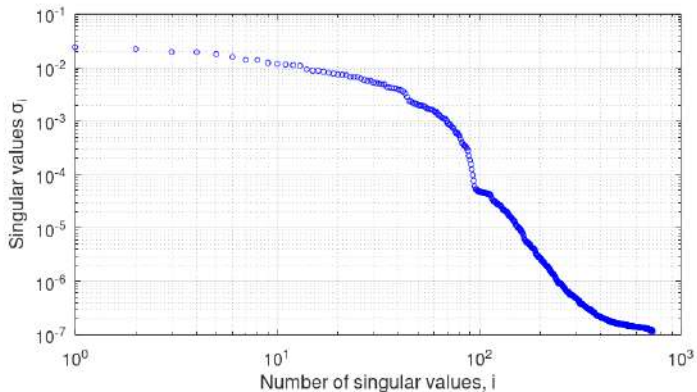


Рис. 3 – Сингулярные значения для сигнала с закрытыми глазами

Компоненты, получаемые методом ССА, являются нормированными (в смысле стандартной евклидовой нормы), вся информация об амплитудах содержится в сингулярных значениях. Для примера на рис. 3 приведены сингулярные значения для сигнала с закрытыми глазами, значения амплитуд первых 10 элементарных компонент на порядок больше остальных. На рис. 4 показаны первые 10 гармоник, видно, что присутствуют гармоники со схожими частотными характеристиками, которые имеет смысл объединять.

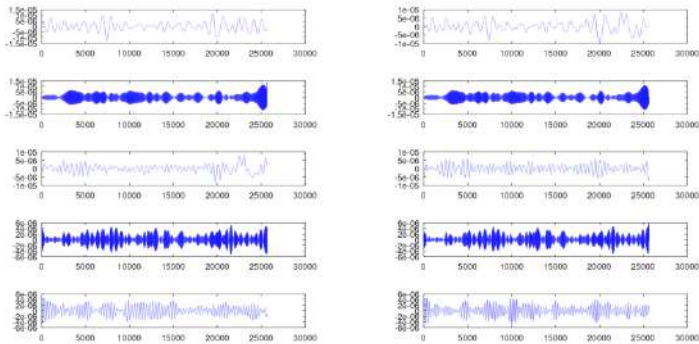


Рис. 4 – Первые 10 элементарных компонент для сигнала с закрытыми глазами.

В заключение необходимо отметить, что предлагаемый подход, включающий сочетание традиционного Фурье-анализа, применения метода ССА и дальнейшей постобработки полученных результатов позволит, на взгляд авторов, повысить точность и однозначность анализа ЭЭГ с целью определения показателей когнитивных процессов деятельности человека.

#### Литература:

1. Berger H. *Über das elektroenzephalogramm des menschen* // arch. F. psychiat. U. nervenkrankh. – 1929. – 87. – S. 527-570.
2. Electrophysiological signatures of resting state networks in the human brain / Mantini D., Perrucci M.G., Del Gratta C., Romani G.L., Corbetta M. // Proc. Natl. Acad. Sci. U S A. -2007. – Vol.104. -№32. – P.: 13170-13175.
3. Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses / Cooper NR, Croft RJ, Dominey SJ, Burgess AP, Gruzelier JH. // Int. J. Psychophysiol. – 2003. – Vol.47. - №1. – P.:65-74
4. EEG alpha oscillations: The inhibition-timing hypothesis / Klimesch W., Sauseng P., Hanslmayr S. // Brain Research Reviews. – 2007. – Vol.53. - №1. – P.:63 – 88.
5. Ливанов М.Н., Гаврилова Н.А., Асланов А.С. Взаимные корреляции между различными участками коры головного мозга при умственной работе // Журн. Выssh. Нерв. Деят. – 1964. – Т. 64, N 2. – С. 185-194..
6. Spatial-temporal structures of human alpha rhythms: theory, microcurrent sources, multiscale measurements, and global binding of networks / Nunez P., Wingeier B., Silberstein R // Hum. Brain Mapp. – 2001. – Vol. 13. – P. 125–164.

7. Intelligence and EEG Phase Reset: A Two Compartmental Model of Phase Shift and Lock / R. W. Thatcher, D. M. North and C. J. Biver // *Neuroimage*. – Vol. 42. – No. 4. – 2008. – P. 1639-1653.

8. Alpha frequency, reaction time, and the speed of processing information / P Klimesch W., Doppelmayr M., Schimke H., Pachinger T. // *J. Clin. Neurophysiol.* — 1996. -Vol.13. - №6. – P. 511-518.

9. Theta synchronization and alpha desynchronization in a memory task / Klimesch W., Doppelmayr M., Schimke H., Ripper B. // *Psychophysiology*. – 1997. – Vol.34. - №2. – P. 169-176.

10. Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis // *Brain Research Reviews*. – 1999. – Vol.29. -№2-3. – P. 169-195.

11. Базанова О.М. Возможность электроэнцефалографических методов диагностики музыкальных способностей // *Биоуправление. Теория и практика*. – 2002. - № 4. – С. 361—365

12. Лебедев А.Н., Мышкин И.Ю., Бовин Б.Г. Оценка психологических параметров личности по электроэнцефалограмме // *Психол. Журн.* — 2002. — Т. 23, № 3. — С. 96-104

13. Базанова О.М. Современная интерпретация альфа-активности ЭЭГ// *Успехи физиологических наук*. – 2009. – No 3. – С. 32.

14. Кулаичев А.П. Компьютерная электрофизиология и функциональная диагностика // *Изд. 4-е, перераб. И доп.* – М.: ИНФРА-М. – 2007. – С. 178-230

15. Начарова М.А., Махин С.А., Павленко В.Б. Оценка реактивности и локализации частотных компонентов альфа-ритма ЭЭГ при выполнении и наблюдении движений // *Рос. Физиол. Журн. Им. И. М. Сеченова*. – 2019. – 105(3). – С. 311-326. – DOI: 10.1134/S0869813919030051

16. Соболев М.Е., Горелик А.Л., Власова О.Л. Разработка и применение нового физико-математического метода анализа количественных электроэнцефалограмм // *Научно-технические ведомости СПбГПУ. Физико-математические науки*. – 2020. – Т. 13. - № 3. – С. 119–129. – DOI: 10.18721/JPM.13309

17. Соловьев А.Ю., Толоконников П.В. Применение метода сингулярно-спектрального анализа в обработке электроэнцефалограмм головного мозга // *Перспективы развития информационных технологий*. – 2014. - № 18. – С. 46-51.

18. Кривецкий А.А. Оценка параметров гармонического сигнала по дискретным наблюдениям в задаче уточнения эфемеридно-временной информации // *Наука. Технологии. Инновации: сб. науч. Тр.* -. – Новосибирск: Изд-во НГТУ, 2022. – С. 114-118.

# МОДЕЛИРОВАНИЕ ЭФЕМЕРИД ПЕРСПЕКТИВНЫХ КОСМИЧЕСКИХ АППАРАТОВ

А.А. Кривецкий

Новосибирский государственный технический университет,  
г. Новосибирск, kriveczkij.2020@stud.nstu.ru

Научный руководитель: Карманов В.С., к.т.н., доцент

*Разработано программное обеспечение для моделирования движения искусственных спутников Земли. Показан результат работы программы для одного из спутников группировки ГЛОНАСС.*

*Software has been developed for simulating the motion of artificial Earth satellites. The program's results for one of the GLONASS satellite constellation members are presented.*

Принцип работы глобальных навигационных спутниковых систем (ГНСС) основан на определении расстояния от объекта (потребителя) до спутников, положение которых должно быть известно с большой точностью. Определение эфемерид космического аппарата (КА) в реальном времени зачастую невозможно [1], так что задача прогнозирования траекторий является актуальной.

Цель данной работы – разработка метода высокоточного моделирования движения КА и исследование перспективных орбитальных группировок. Основная сложность при моделировании движения на околоземной орбите – это вычисление действующих на КА сил. В данной работе рассмотрено влияние Земли (с учетом геопотенциалов), Луны и Солнца.

Для околоземной орбиты используются системы координат с началом отсчета в центре Земли.

## **Системы координат:**

1. *Земная система.* Жёстко связана с телом Земли и вращается вместе с ней.
  - 1.1. *Истинная экваториальная.* Основная плоскость – истинный экватор. Ось абсцисс направлена в истинную точку весеннего равноденствия.
  - 1.2. *Средняя экваториальная.* Основная плоскость – средний подвижный экватор. Ось абсцисс направлена в среднюю точку весеннего равноденствия.
2. *Небесная система.* Задана положением экватора, фиксированного на стандартную эпоху J2000.0. Ось абсцисс направлена в среднюю точку весеннего равноденствия, соответствующую стандартной эпохе.

Преобразование из небесной системы координат ( $C$ ) в земную ( $T$ ) выполняется по формуле:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_T = R_2(-x_p) \cdot R_1(-y_p) \cdot R \cdot N \cdot P \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}_C, \quad (1)$$

где  $P$ - матрица прецессии,  $N$  – матрица нутации,  $R = R_3(S_{\oplus})$  – матрица вращения Земли,  $R_1(-y_p)$  – матрица поворота вокруг оси абсцисс против часовой стрелки на эмпирическое значение координаты полюса Земли  $y_p$ ,  $R_2(-x_p)$  - матрица поворота вокруг оси ординат против часовой стрелки на эмпирическое значение координаты полюса Земли  $x_p$ . Матрицы прецессии  $P$  и нутации  $N$  вычисляются с помощью углов прецессии и нутации соответственно. Сами углы вычисляются по эмпирическому закону [2]. Аргументом матрицы вращения Земли является гринвичское истинное звёздное время  $S_{\oplus}$  - часовой угол истинной точки весеннего равноденствия.

Обратное преобразование осуществляется путем транспонирования матрицы перехода.

#### **Временные шкалы:**

1. Всемирное координированное время (UTC).
2. Всемирное время (UT1).
3. Равномерная шкала земного времени (TT).
4. Бариецентрическое динамическое время (TDB).

Моменты времени в различных шкалах выражаются в модифицированных юлианских днях (MJD).

В пределах солнечной системы пользуются равномерной шкалой бариецентрического времени. В этой шкале вычисляются положения Луны и Солнца. Для перевода из шкалы земного времени в модели использована модель, составленная NASA [3].

Всемирное время является суммой всемирного координированного времени и поправки  $dUT$ , которая может быть найдена только с помощью наблюдений. Поправка  $dUT$  вместе с координатами полюса  $x_p$  и  $y_p$  (1), составляют параметры вращения Земли [4-5].

Более подробное описание систем координат и временных шкал приводится в литературе [6].

Геопотенциал в земной системе координат имеет вид:

$$U = \frac{fm}{r} \left\{ 1 + \sum_{n=2}^{N_{max}} \sum_{k=0}^n \left( \frac{r_0}{r} \right) P_n^{(k)}(\sin \varphi) [C_{nk} \cos k\lambda + S_{nk} \sin k\lambda] \right\}, \quad (2)$$

где  $fm, r_0$ -геоцентрическая гравитационная постоянная и экваториальный радиус Земли,  $C_{nk}, S_{nk}$  – числовые коэффициенты при зональных, тессеральных и секториальных гармониках разложения гравитационного поля Земли,  $P_n^{(k)}$  - присоединенные функции Лежандра.



Приведенная модель гравитационного поля Земли и числовые значения коэффициентов приведены в работе [7].

Солнце и Луна оказывают возмущения на движение КА, которые нельзя игнорировать при высокоточном моделировании орбиты. Полное моделирование движения Солнца и Луны относительно Земли подразумевает моделирование движения всех объектов Солнечной системы (задача N-тел), что является очень ресурсоемким процессом. Однако, можно воспользоваться упрощенной моделью, основанной на наблюдениях, проведенных с Земли [8, 9].

Положения вычисляются в эклиптической (сферической) системе координат. Для вычисления ускорений, действующих на КА необходимо переводить эклиптические координаты в небесные. Переход производится по следующим формулам:

$$\begin{aligned}x &= r \cos \delta \cos \alpha, \\y &= r \cos \delta \sin \alpha, \\z &= r \sin \delta,\end{aligned}$$

где  $\alpha$  – долгота [рад],  $\delta$  – широта [рад],  $r$  – расстояние.

Для вычисления положений Луны и Солнца используются фундаментальные аргументы:  $l$  – средняя аномалия Луны,  $l'$  – средняя аномалия Солнца,  $F$  – средний аргумент широты Луны,  $D$  – разность средних долгот Луны и Солнца,  $\Omega$  – средняя долгота восходящего узла Луны. Для вычисления положения Солнца необходимо учитывать средние аномалии Венеры, Земли, Марса, Юпитера и Сатурна.

Пусть  $\vec{r}$ ,  $\vec{v}$  – векторы положения и скорости КА относительно центра Земли в системе экватора и эклиптики, фиксированной на стандартную эпоху J2000.0 (небесная система координат). Тогда уравнение движения примет вид:

$$\frac{d^2\vec{r}}{dt^2} = \vec{F}_E + \vec{F}_M + \vec{F}_S + \vec{F}_a, \quad (4)$$

где  $\vec{F}_E$  – ускорение, обусловленное геопотенциалом,  $\vec{F}_M$  – ускорение, вызываемое Луной,  $\vec{F}_S$  – ускорение, вызываемое Солнцем,  $\vec{F}_a$  – ускорение, обусловленное торможением в атмосфере.

Общепризнанным методом для вычисления орбит [10] является неявный одношаговый метод численного интегрирования, разработанный Э. Эверхартом [11]. Суть метода – представление второй производной функции в виде полинома  $N$ -ой степени. В данной работе используется 7 степень.

Разработанное программное обеспечение позволяет моделировать движение искусственных спутников Земли, как существующих, так и

перспективных для ГНСС. Пример моделирования для спутника ГЛОНАСС-R01 в период с 19.05.2023 по 20.05.2023 представлен на рис. 1.

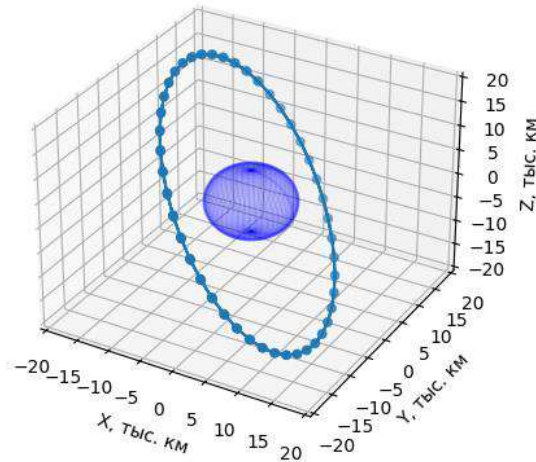


Рис. 1 - Пример суточного вращения спутника ГЛОНАСС (2 витка).

Таким образом, на данном этапе разработки вычисляется орбита КА с любой дискретностью. Дальнейшим этапом необходимо оценивать точность геопозиционирования. Это позволит находить оптимальные в различных смыслах структуры орбитальной группировки.

#### Литература:

1. Современные и перспективные информационные ГНСС технологии в задачах высокоточной навигации / Под ред. В.А. Бартечева, М.Н. Красильщикова. – М.: Физматлит, 2014. – 192 с.
2. Абалакин В.К., Аксёнов Е.П., Гребеников Е.А., Дёмин В.Г., Рябов Ю.А. Справочное руководство по небесной механике и астродинамике. Под ред. Г.Н.Дубошина. М., Наука, 1976
3. POLYNOMIAL EXPRESSIONS FOR DELTA T [Электронный ресурс] // nasa.gov. Сайт национального управления по авиации и исследованию космического пространства (NASA, США) URL: <https://eclipse.gsfc.nasa.gov/SEhelp/deltatpoly2004.html>
4. Кривецкий, А. А. Прогнозирование параметров вращения и движения земли методом SSA / А. А. Кривецкий // Дни науки НГТУ-2022: МАТЕРИАЛЫ НАУЧНОЙ СТУДЕНЧЕСКОЙ КОНФЕРЕНЦИИ (ИТОГИ НАУЧНОЙ РАБОТЫ СТУДЕНТОВ ЗА 2021-2022 ГГ.),

Новосибирск, 14 марта – 18 2022 года. – Новосибирск: Новосибирский государственный технический университет, 2022. – С. 95-98.

5. Модификация базового метода сингулярного спектрального анализа для повышения точности прогнозирования неравномерности вращения Земли / А. К. Гречкосеев, А. С. Толстиков, В. М. Тиссен [и др.] // Вычислительные технологии. – 2020. – Т. 25, № 3. – С. 54-65.

6. Куимов К.В. Редукционные вычисления. / В сб.: Практикум по астрометрии. изд.-во Московского университета, 1989, с.6-42.

7. Tapley B.D., Watkins M.M., Ries J.C., Davis G.W., Eanes R.J., Poole S.R., Rim H.J., Schutz B.E., Shum C.K., Nerem R.S., Lerch F.J., Marshall J.A., Klosko S.M., Pavlis N.K., and Williamson R.G. The Joint Gravity Model 3. //Journ. Geophys. Res., 1996, v.101, pp.28029-28049.

8. Абалакин В.К., Аксёнов Е.П., Гребеников Е.А., Дёмин В.Г., Рябов Ю.А. Справочное руководство по небесной механике и астродинамике. Под ред. Г.Н.Дубошина. М., Наука, 1976.

9. Эфемеридная астрономия. / Труды ИПА РАН, 2004, т.10.

10. Бордовицына Т.В. Современные численные методы в задачах небесной механики. М., Наука, 1984.

11. Everhart E. Implicit single-sequence methods for integrating orbits. // Celestial Mechanics, 1974, vol.10, p.35-55.

## **ПРИМЕНЕНИЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ КЛАССИФИКАЦИИ ДВИЖЕНИЙ НА ОСНОВЕ ЭЛЕКТРОМИОГРАММ**

**И.В. Кузнецов**

**Дальневосточный государственный университет путей сообщения,  
г. Хабаровск, [ivankuznetsov@festu.khv.ru](mailto:ivankuznetsov@festu.khv.ru)**

*При классификации движений на основе сигнала электромиограммы большую роль играет его предобработка. Ставится вопрос об эффективности применения вейвлет-преобразований для ЭМГ-сигнала с целью дальнейшей классификации движений. Для проверки гипотезы проведен эксперимент по распознаванию движений с помощью классификатора – многослойного перцептрона. Результаты эксперимента позволяют утверждать о высокой эффективности вейвлетов Добеши, а также о большом потенциале составления ансамбля из вейвлет-преобразования и быстрого преобразования Фурье.*

*Preprocessing of the electromyogram signal plays an important role in solving the problem of classifying movements based on it. This paper is devoted to the problem of the effectiveness of using wavelet transforms for the EMG-signal in order to further classification of movements. An experiment was conducted on motion recognition using*

*multilayer perceptron to evaluate the effectiveness of using wavelet transforms. The research results suggest the good efficiency of Daubechies Wavelets, as well as the great potential of composing an ensemble from the wavelet transform and the Fast Fourier Transform.*

Одним из перспективных методов, способных лечь в основу человеко-машинного интерфейса (ЧМИ) является электромиография (ЭМГ). Электромиография – это метод регистрации биоэлектрических потенциалов, возникающих при напряжении мышечных волокон.

Многими исследователями предлагается использовать поверхностную электромиографию (пЭМГ) для классификации движений руки. Решение данной задачи может позволить создать жестовый интерфейс, не накладывающий для своей работы иных условий кроме необходимости получения сигналов электромиографии (ЭМГ-сигналов). Так, в работе [1] предложена стратегия управления роботизированными устройствами, осуществляющая совмещенное командно-пропорциональное управление на основе сигналов нейромышечной активности. В качестве устройства-считывателя использовался браслет Thalmic Labs Myo. Авторы [2] указывают, что имеется возможность применения подхода к классификации данных, основанного на базовых и ансамблевых методах машинного обучения и приводят результаты применения методики испытаний программ для классификации паттернов сигналов электромиографии с использованием методов «случайного леса», «AdaBoost» и «CatBoost». В работах [3,4] рассмотрены основные признаки, которые могут быть извлечены из ЭМГ-сигнала, а также приведено описание наиболее часто используемых классификаторов. Авторы указывают, процесс выбора метода извлечения признаков очень субъективен, однако многие методы, основанные на временной области, отображают результаты аналогичной формы.

Особый интерес в исследованиях ЭМГ-сигналов для решения задачи классификации движений является их предобработка и извлечение признаков. При этом часто применяются Быстрое преобразование Фурье (БПФ). Кроме этого, во многих исследованиях нашли применение вейвлет-преобразования. Авторы [5] провели анализ ЭМГ-сигнала с использованием распределения временных частот (TFD), представляющего собой спектрограмму с различным размером окна. В работе указывается, что при БПФ не дает возможности работы с временным представлением сигнала, с другой же стороны вейвлет-преобразования показывают большую вычислительную сложность. Однако вопрос об эффективности классификации движений на основе сигнала, подвергнутого вейвлет-преобразованию, остается открытым.

Для рассмотрения данной проблемы были построены классификаторы на основе алгоритма многослойного персептрона (MLP). Классификаторы обучались на наборе данных, представленном в работе [6]. Каждый классификатор обучался в двух вариантах: на основе данных без применения дополнительных преобразований, и на основе данных, подвергнутых вейвлет-преобразованию с использованием вейвлета Добеши.

Для классификации данных были выбраны следующие признаки:

- максимальное значение сигнала;
- минимальное значение сигнала;
- максимальное абсолютное значение;
- стандартное отклонение;
- среднее абсолютное значение (MAV):

$$MAV = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (1)$$

– длина волны (WL) – является мерой сложности сигнала и показывает длину волны в пределах временного окна:

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (2)$$

– число пересечений сигналом нулевого значения (ZC):

$$ZC = \sum_{i=1}^{N-1} \text{sgn}(x_{i+1} * x_i) \cap |x_i - x_{i+1}| \geq \text{threshold}, \quad (3)$$

где *threshold* – пороговое значение, которое используется для нивелирования влияния шумов,

$$\text{sgn} = \begin{cases} 1, & \text{если } x \geq \text{threshold} \\ 0, & \text{если } x < \text{threshold} \end{cases}$$

– среднее квадратическое (RMS):

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (4)$$

– среднее изменение амплитуды (AAC):

$$AAC = \frac{1}{N} \sum_{i=1}^N |x_{i+1} - x_i| \quad (5)$$

– амплитуда Виллисона (WAMP):

$$WAMP = \sum_{i=1}^{N-1} f(|x_i - x_{i+1}|) \quad (6)$$

$$f(x) = \begin{cases} 1, & \text{если } x \geq \text{threshold} \\ 0, & \text{если } x < \text{threshold} \end{cases}$$

Для реализации многослойного персептрона использованы язык программирования Python версии 3.10 и фреймворк Tensorflow версии 2.14.0.

В случае использования набора данных, не подвергнутого вейвлет-преобразованию, метрики Accuracy, Precision и Recall классификатора приобретают вид, показанный на рисунке 1. Из графиков видно, что при высоких достигаемых значениях Accuracy и Recall, эффективность классификатора с точки зрения метрики Precision невелика. Это говорит

о том, что в случае использования необработанных данных затрудняется задача различения классов движений друг от друга. Это может быть обусловлено как похожестью различных движений с точки зрения затрагиваемых мышц, так и влиянием шумов и интерференции электромиограмм многих мышц на итоговую форму рассматриваемого сигнала.

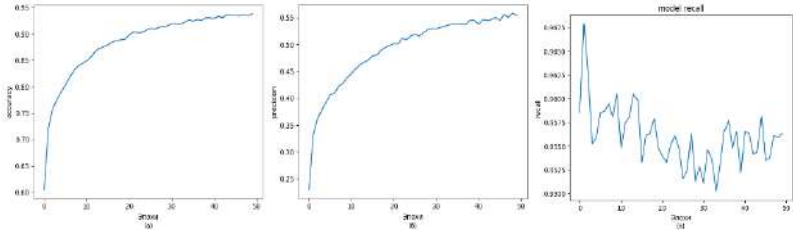


Рис.1 – метрики эффективности классификатора при необработанных данных: а – accuracy, б – precision, в – recall

В случае же использования данных, обработанных с помощью вейвлет-преобразования, метрики эффективности работы классификатора примут вид, показанный на рисунке 2. Графики показывают улучшение работы классификатора как с точки зрения метрики Recall, так и с точки зрения метрики Precision, что говорит о том, что применение вейвлет-преобразования к исходным данным улучшает различимость различных классов движений.

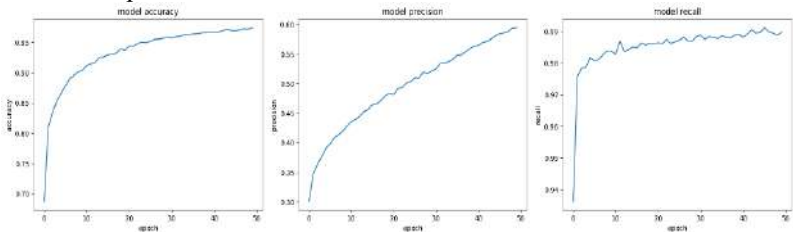


Рис.2 – метрики эффективности классификатора при данных с использованием вейвлет-преобразования: а – accuracy, б – precision, в – recall

Результаты эксперимента показывают, что предварительная обработка данных играет важную роль в решении задачи классификации движений на базе ЭМГ-сигнала. Ожидается, что выбор вейвлет-преобразования и его параметров может позволить дополнительно повысить эффективность работы классификатора. Следует также отметить, что при использовании вейвлет-преобразования не затрагивались признаки частотной области. Перспективным направлением кажется совместное применение вейвлет-преобразований

и быстрого преобразования Фурье для дальнейшей модификации работы классификатора.

Литература:

1. Лобов С.А., В.И. Миронов, И.А. Кастальский, В.Б. Казанцев Совместное использование командного и пропорционального управления внешними робототехническими устройствами на основе электромиографических сигналов // Соврем. технол. мед.. 2015. №4. URL: <https://cyberleninka.ru/article/n/sovmestnoe-ispolzovanie-komandnogo-i-proporsionalnogo-upravleniya-vneshnimi-robototekhnicheskimi-ustroystvami-na-osnove> (дата обращения: 08.11.2023).
2. Методика испытаний программ для классификации паттернов сигналов электромиографии / А. А. Кабанов, Г. В. Никонова, А. В. Никонов, Р. Н. Богатов // Информационные технологии и автоматизация управления : Материалы XIV Всероссийской научно-практической конференции студентов, аспирантов, работников образования и промышленности, Омск, 26–27 мая 2023 года / Отв. редактор А.В. Никонов. – Омск: Омский государственный технический университет, 2023. – С. 104-116. – EDN MEENYM.
3. Christopher S., MdRasedul I. A Comprehensive Study on EMG Feature Extraction and Classifiers // Open Access Journal of Biomedical Engineering and Biosciences 2018. № 1 (1). URL: <https://lupinepublishers.com/biomedical-sciences-journal/pdf/OAJBEV.MS.ID.000104.pdf> (дата обращения: 08.11.2023).
4. Phinyomark A., Phukpattaranont P., Limsakul C.. Feature reduction and selection for EMG signal classification // Expert Systems with Applications 2012. Т. 39. № 8. С. 7420-7431.
5. Electromyography Signal Analysis Using Spectrogram / T. N. S. T. Zawawi, A. R. Abdullah, E. F. Shair, I.Halim, Rawaida.O // 2013 IEEE Student Conference on Research and Development (SCORED) 2013. С. 319–324.
6. EMG Based Control of Individual Fingers of Robotic Hand / Naseer N., Faizan A., Sameer A., Saad I., Rayyan A. K., Hammad N. // 2018 International Conference on Sustainable Information Engineering and Technology (SIET). 2018 № 1. С. 6–9..

## **ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОЦЕНКИ СТОИМОСТИ ИМУЩЕСТВА**

**О.В. Курочкин**

**Новосибирский государственный университет  
экономики и управления,**

**г. Новосибирск, kurochkinoleg228@gmail.com**

**Научный руководитель: Семенов Н.С., к.э.н., доцент**

*В статье рассматриваются вопросы применения искусственного интеллекта для оценки стоимости имущества. Выявлены положительные стороны применения искусственного интеллекта в данной сфере, а также возможные негативные последствия.*

*The article discusses the use of artificial intelligence to assess the value of property. The positive aspects of the use of artificial intelligence in this area, as well as possible negative consequences, have been identified.*

Актуальность выбранной темы обусловлена тем, что стремительное развитие информационных технологий, в том числе, искусственного интеллекта (далее ИИ), приводит к возникновению множества вопросов, как технических, так и философских, правовых, организационных, связанных с использованием ИИ в различных сферах человеческой деятельности. Сегодня искусственный интеллект решает ряд задач, которые ранее решались человеком, вытесняя с рынка определенные профессии (дизайнер, копирайтер, программист (пока невысокой квалификации)). Современные тенденции развития ИИ показывают его возможности в проведении оценки стоимости объектов недвижимости (Сбербанк РФ внедряет эксперимент по использованию нейросети в процессе оценки недвижимости), что приводит к возникновению сразу нескольких вопросов: 1) каковы возможности и перспективы развития искусственного интеллекта в сфере оценки стоимости имущества? 2) вытеснит ли ИИ оценщиков, как постепенно вытесняет дизайнеров и копирайтеров?

Попробуем рассмотреть особенности использования ИИ при оценке недвижимого имущества, отметим,

Следует отметить, что рынок недвижимости, наряду с финансовым рынком – это важнейшая составляющая мировой экономики, соответственно, на функционирование и развитие данного рынка направлено множество различных ресурсов, которые ориентированы, в том числе, на оптимизацию работы субъектов рынка.

Рынок искусственного интеллекта оценивается более чем в 300 миллиардов долларов США с перспективой увеличения до 1 триллиона



долларов США в 2029 году. Эта ценность является результатом широкого использования искусственного интеллекта в финансовом секторе, а также в сфере недвижимости. Именно возможности анализа рынка недвижимости, выявления им тенденций и определения стоимости, например, при принятии решения о выдаче ипотечного кредита обусловили его развитие в сфере оценки недвижимости.

Попробуем обозначить возможности и перспективы развития искусственного интеллекта в сфере оценки стоимости имущества (в том числе – недвижимого):

Во-первых, возможность существенно облегчить работу оценщиков в сфере сбора и анализа информации, поиска аналогов, систематизации и обработки информации. Возможности обработки огромных массивов данных выступает важным фактором, привлекающим к ИИ специалистов рынка недвижимости, оценщиков. Экономия времени на анализ и оценку составляет, по данным специалистов Сбербанка РФ, внедряющих технологию ИИ, до 50% [3].

Во-вторых, возможности прогнозирования развития рынка и цен на основе поведения акторов рынка. Искусственный интеллект используется для сопоставления тысяч переменных, связанных с жильем — ставок по ипотечным кредитам, качества школ, количества спален, занятости на местном уровне — для построения моделей, прогнозирующих цены на жилищные активы.

В-третьих, возможность оценивать и прогнозировать риски, связанные с оценочной деятельностью, влияющие на стоимость объекта, в том числе те, которые не всегда может учесть человек.

В-четвертых, беспристрастность и объективность: ИИ не испытывает эмоций и действует строго в рамках алгоритма, что снижает риски как ошибок, так и злоупотреблений.

Далее представим возможные ограничения применения ИИ в оценке стоимости имущества:

1) ИИ не всегда способен учесть при стоимости объекта нестандартные условия, поскольку ограничен заложенными алгоритмами;

2) зависимость от достоверности, доступности поступающей информации, находящейся в открытом доступе, также влияет на деятельность по оценке, соответственно, существуют различия в оценке одного и того же объекта разными электронными инструментами, что требует коррекции человеком, берущим на себя ответственность за результаты оценки;

3) нерешенность ряда правовых и организационных вопросов применения искусственного интеллекта – сегодня именно оценщик и

оценочная организация несут ответственность за результаты оценки своей репутацией, финансовыми обязательствам, что не может делать искусственный интеллект. Поэтому может возникнуть ряд правовых споров, связанных с правовым статусом и возможностью принятия решений о стоимости квартиры или другого объекта на основе мнения ИИ.

Достаточно сложно сказать, каким образом будет проходить развитие ИИ в оценочной деятельности, и какую долю рынка займет он на рынке оценочных услуг. Сегодня инвестиции в развитие ИИ в мире велики: в 2021 году они составили 125,4 млрд. долл., а в 2022 году – 91,9 млрд. долл. На наш взгляд, искусственный интеллект сегодня обеспечивает возможности для облегчения труда оценщиков, возможно, сокращение трудоемкости приведет к сокращению числа оценщиков на 10-15%, но говорить о каких-либо конкретных цифрах сокращения доля рынка людей-оценщиков достаточно сложно. Сегодня рынок оценки испытывает определенный кадровый голод, возможности ИИ в данном случае позволят снизить негативные последствия от этого явления, возможно, появится возможность для привлечения к процессам обработки данных сотрудников с более низкой квалификацией, а деятельность высококвалифицированных оценщиков будет сводиться к проверке, подписанию отчетов.

Таким образом, можно сделать вывод о том, что ИИ, безусловно, является эффективным инструментом для обработки, анализа и представления информации о том или ином объекте, однако, на наш взгляд, мнение оценщика крайне важно для подтверждения данных, решения вопросов полномочий и ответственности, поэтому ИИ сегодня, хоть и способен облегчить работу оценщиков, не должен заменять их при принятии итогового решения.

#### Литература:

1. *Выходцев Н.А.* Использование искусственного интеллекта для оценки стоимости недвижимого имущества /Н.А. Выходцев// Доклады Томского государственного университета систем управления и радиоэлектроники. – 2021. – №1. – С. 68-72.

2. *Чулюкова С.А., Чулюкова Н.А.* Искусственный интеллект и креативное мышление в оценке объектов недвижимости в российской федерации //Юридическая гносеология. – 2022. – № 4. – С. 92-99.

3. *Красовский К.А.* Внедрение технологий искусственного интеллекта в кадастровую оценку недвижимости в России //Современные проблемы землеустройства, кадастров, природообустройства и повышения безопасности труда в АПК. – Красноярск, 2023. – С. 187-190.

4. Ласкин М.Б., Гадасина Л.В. Особенности применения методов, основанных на деревьях решений, в задачах оценки недвижимого имущества //Бизнес-информатика. – 2022. – Т.16. – № 4. – С. 7-18.

## **ПРОГНОЗИРОВАНИЕ ЭФФЕКТИВНЫХ СВОЙСТВ ОДНОНАПРАВЛЕННЫХ КОМПОЗИТНЫХ МАТЕРИАЛОВ С ПОМОЩЬЮ ОДНОСЛОЙНОЙ НЕЙРОННОЙ СЕТИ**

**А.П. Никишин**

**Санкт-Петербургский политехнический  
университет Петра Великого,  
г. Санкт-Петербург, [nikishin.ap@edu.spbstu.ru](mailto:nikishin.ap@edu.spbstu.ru)  
Научный руководитель: Гриценко А.И.**

*Композиционные материалы с однонаправленной структурой находят широкое применение в различных отраслях: гражданское и промышленное строительство, инфраструктура нефтегазовой промышленности. Их выдающиеся механические свойства делают их важными компонентами в создании прочных и легких конструкций. Однако, при разработке и применении таких композитных материалов возникают сложности, связанные с определением их эффективных свойств, эксперименты обычно очень сложные и дорогостоящие.*

*Composite materials with unidirectional structures find broad applications in various industries: civil and industrial construction, infrastructure in the oil and gas sector. Their outstanding mechanical properties make them essential components in the creation of strong and lightweight structures. However, when developing and utilizing such composite materials, challenges arise in determining their effective properties, as experiments are typically complex and costly.*

Целью работы является разработка методов оценки эффективных упругих свойств однонаправленных полимерных волокнистых композиционных материалов, основанных на применении нейросетевых технологий.

Механические свойства композиционных материалов на макроуровне прежде всего зависят от свойств отдельных компонент. В данной работе рассматривается задача прогнозирования эффективных упругих свойств полимерного однонаправленного волокнистого композиционного материала в рамках предположения о трансверсальной изотропности его свойств. В этом случае тензор упругих податливостей имеет 5 независимых компонент:  $E_{11}$ ,  $E_{22}$ ,  $G_{12}$ ,  $G_{23}$ ,  $\nu_{12}$ .

Материал матрицы и включения принимается изотропными, имеющими 2 независимых компоненты:  $E$ ,  $\nu$

Решение задачи гомогенизации может быть осуществлена как на основе аналитических оценок, так и при помощи метода конечно элементной (КЭ) гомогенизации [1]. Среди аналитических оценок можно выделить модели А. Рейса [2] или В. Фойгта [2], которые легли в основу правила смесей, другие подходы базируются на решениях задачи Эшелби [3,4] об эллиптическом включении в бесконечную однородную упругую среду.

Основным недостатком аналитических оценок следует считать их низкую точность для композитов со сложной топологией. Применение метода КЭ гомогенизации позволяет напрямую учесть все особенности микроструктурной организации композитных материалов, однако, ценой этому становится высокая вычислительная сложность. В последнее время широкое применение для оценки макросвойств гетерогенных материалов получили методы машинного обучения [6].

В работе рассматривалась простейшая полносвязная нейронная сеть с одним скрытым слоем, на вход которой подаются свойства матрицы и включения и доля включения (Рисунок 1).

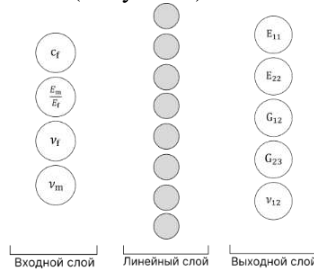


Рис. 1 – Схематичное изображение нейронной сети

Для создания обучающей выборки для однослойной нейронной сети рассматривались различные правила оценок свойств для трансверсально-изотропного материала: правило смеси (1), модель Халпина-Цая, модель Хашина-Розена [2] и модель Мори-Танака [3,4].

$$E_{11} = (1 - c_f)E_m + c_f E_f; E_{22} = \frac{E_f E_m}{(1 - c_f)E_f + c_f E_m} \quad (1)$$

Для оценки точности аналитических решений, полученные результаты сравниваются с результатами, полученными методами КЭ гомогенизации по формуле:

$$Error(E_{11}) = \frac{|E_{11}^{МКЭ} - E_{11}^A|}{E_{11}^{МКЭ}} \cdot 100\%, \quad (2)$$

где  $E_{11}^{МКЭ}$  – эффективный модуль Юнга в продольном направлении для волокна, полученный с помощью МКЭ гомогенизации,  $E_{11}^A$  – результат, полученный одним из аналитических способов.

Был проведен анализ влияния объемной доли включения на эффективные свойства изотропного композитного материала, имеющего свойства:  $E_f = 72$  ГПа,  $\nu_f = 0.36$ ,  $E_m = 3.79$  ГПа,  $\nu_m = 0.38$ . Результаты проведенных вычислительных экспериментов показали, что изменение объемной доли существенно влияет на точность оценок (Рисунок 2). Поэтому при составлении обучающей выборки также необходимо учитывать данный параметр модели.

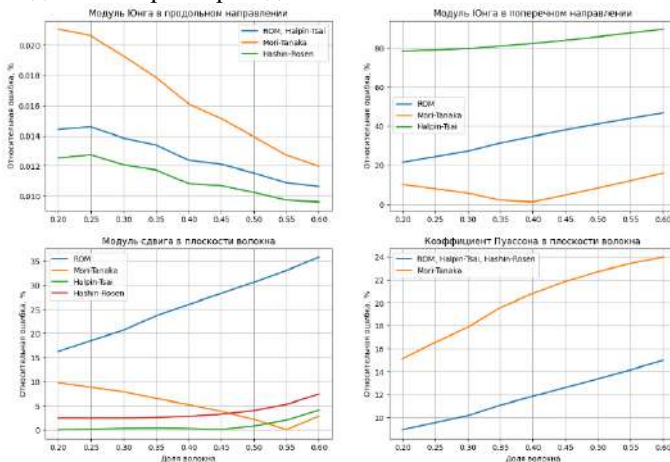


Рис. 2 – Влияние объемной доли включения

На основе анализа результатов о точности предсказания различных аналитических оценок, была подготовлена обучающая выборка, отдельные компоненты которой соответствуют моделям, показавшим лучший, с точки зрения точности, результат. Полученный набор данных содержит более 300 тысяч значений и служит основой для дальнейших исследований и анализа.

В данном эксперименте была проведена оценка влияния нейронов входного слоя, количества нейронов в линейном слое и различных функций активации на функцию потерь в процессе обучения нейронной сети. Для оценки качества обучения использовалась функция потерь MSE, были построены графики зависимости Average MSE (усреднение ошибки по всей валидируемой выборке) от цикла обучения для различных параметров моделей (Рисунок 3).

Оптимизация обучения нейронной сети выполнялась с использованием алгоритма Adam, который комбинирует два момента для коррекции шага обновления весов: скользящее среднее градиентов и квадратов градиентов.

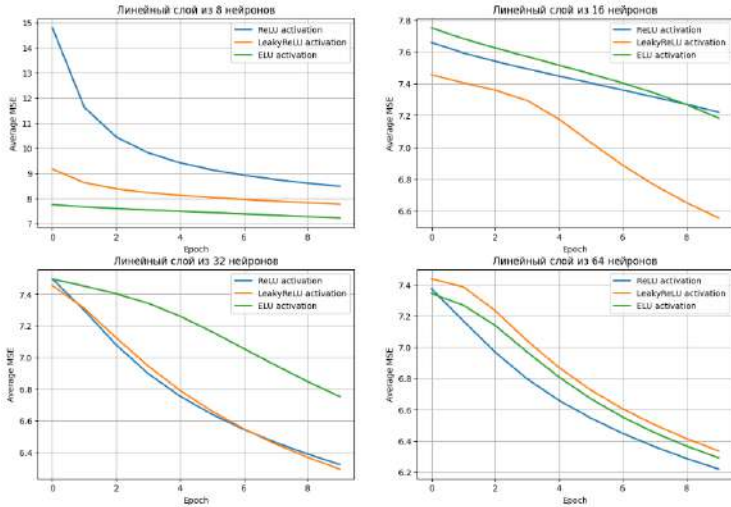


Рис. 3 – Результаты обучения нейронной сети

С целью верификации предложенных методик было проведено сравнение результатов (Таблица 1) аналитических методов, МКЭ и модели нейронной сети с экспериментальными данными  $E_{11}^{\exists} = 18$  ГПа,  $E_{22}^{\exists} = 6$  ГПа,  $G_{12}^{\exists} = 2$  ГПа [5]. Скрытый слой состоит из 64 нейронов с функцией активации ReLU.

Таблица 1 – Сводная таблица результатов

Метод	$\frac{ E_{11}^{\exists} - E_{11}^i }{E_{11}^{\exists}} \cdot 100,$	$\frac{ E_{22}^{\exists} - E_{22}^i }{E_{22}^{\exists}} \cdot 100,$	$\frac{ G_{12}^{\exists} - G_{12}^i }{G_{12}^{\exists}} \cdot 100,$
МКЭ	0.6331	2.2108	0.7598
Правило смесей	0.6339	21.1456	14.2715
Халпин-Цай	0.6339	7.3924	13.9937
Хашин-Розен	0.6358	-----	0.1759
Мори-Танака	0.6272	10.5216	12.3912
Нейронная сеть	2.0261	0.1926	9.5197

В заключении, была рассмотрена модель однослойной полносвязной нейронной сети для определения эффективных упругих свойств однонаправленных волокнистых полимерных композиционных материалов. Можно свидетельствовать о достижении хорошей точности,

сравнивая с экспериментальными результатами даже при использовании одного линейного слоя с нелинейной функцией активацией.

Литература:

1. Грищенко А. И., Игнатович И. А., Петросян О. Л. Оптимизация микроструктуры композиционных материалов с учетом ограничений на их свойства // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 3. С. 73–86. DOI: <https://doi.org/10.18721/JPM.16307>

2. Муйземнек А.Ю., Иванова Т.Н., Карташова Е.Д. Сопоставление результатов экспериментального и расчетного определения эффективных характеристики упругих свойств полимерных слоистых композитов из угле- и стеклотканей // Вестник Пермского национального исследовательского политехнического университета. Механика. – 2021. – № 2. С. 88–105.

3. Фрейдин А. Б. Задача Эшелби: учеб. пособие / А. Б. Фрейдин. — СПб. : ПОЛИТЕХ-ПРЕСС, 2022. — 148 с

4. Ling Liu. A Note on Mori-Tanaka’s method // Acta Mechanica Solida Sinica. – 2014. – P. 234–244. –doi:10.1016/S0894-9166(14)60033-1

5. Wang H.W., Zhou H.W., Gui L.L., Ji H.W., Zhang X.C.. Analysis of effect of fiber orientation on Young’s modulus for unidirectional fiber reinforced composites // Composites Part B-engineering. – 2014. – Vol. 56. P. 733-739.

6. Rao C., Liu Y., Three-dimensional convolutional neural network (3d-cnn) for heterogeneous material homogenization // Computational Materials Science 184. –2020.

## **РАСПОЗНАВАНИЕ ГОЛОСОВЫХ КОМАНД ПОЛЬЗОВАТЕЛЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ «УМНЫЙ ДОМ» С ИСПОЛЬЗОВАНИЕМ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ**

**Н.В. Патрушев**

**Новосибирский государственный технический университет,  
г. Новосибирск, [patrushev.nikita2@mail.ru](mailto:patrushev.nikita2@mail.ru)**

**Научный руководитель: Авдеенко Т.В., д.т.н., профессор**

*Работа посвящена исследованию и программной реализации модуля для распознавания голосовых команд пользователя. Модуль предназначен для управления системой “умный дом” с помощью пользовательского интерфейса.*

*The work is devoted to the research and software implementation of a module for selecting several user commands. The module is designed to control the smart home system using a user interface.*

В настоящее время существует возможность управлять своим домом на расстоянии. В этом могут помочь компьютеры, пульта дистанционного управления и мобильные телефоны. Лучший вариант — это мобильный гаджет, который всегда под рукой. Голосовое управление умным домом работает по четкому алгоритму, разработанному специально для дома. Управление домом представляет собой серию команд, где ключевые задачи задаются в виде продуманных последовательностей определенных слов. Нейросетевая модель просто распознает определенные звуки. В системе уже запрограммированы некоторые типичные команды.

Целью настоящей работы является исследование нейросетевой модели для распознавания голосовых команд и разработка пользовательского интерфейса для взаимодействия с модулем. Данные для исследования были взяты из датасета Yandex Cup от компании Yandex. Датасет состоит из обучающей и тестирующей выборки. Обучающая выборка состоит из более чем 100 тысяч аудиоданных и 39 категорий голосовых команд длительностью около 3 секунд.

Для обучения подсистемы нейронной сети было выбрано четыре голосовые команды: Дизлайк, Пауза, Замолчи, Вверх. Для нормализации данных для каждой голосовой команды было взято по тысяче элементов голосовых файлов. Нормализация данных подразумевает преобразование аудиосигналов к одному типу. В данной работе был использован следующий алгоритм для нормализации данных:

1. Сжатие данных – изменение стереосигнала на моносигнал.
2. Избавление от задней оси. Из полученного тензора равным 13200 на 1, где 1 – количество каналов, удаляем размерность единицы из формы тензора. Таким образом, получаем вектор равный 13200.
3. Приведение частоты дискретизации к целому типу (int64).
4. Передискретизация сигнала, получение из 44100 Гц частоты 16000 Гц. Служит для уменьшения размера окончательного сигнала.
5. Обрезка сигнала до длины 48000 шагов.
6. Заполнение нулями данных, с количеством шагов меньше 48000.

Зная, что на 1000 элементов приходится 132 аудиоданных с длиной 64000 шагов. Так как голосовая команда может находиться в начале, середине или конце аудиофайла, было принято решение об уменьшении аудиофайла до 48000 шагов. Вероятность получения пустого тензора составляет 33%. Таким образом, вероятность потерять данные составляет около 4%.



Для получения характеристик аудиосигнала был использован метод MFCC (мел-частотные кепстральные коэффициенты), которые вычисляются согласно графику зависимости высоты звука от частоты колебаний [1].

Расчет коэффициентов проводится в несколько шагов:

1. Вычисляются дискретные преобразования Фурье ( $N$  – количество значений сигнала, измененных за период,  $f[k]$  – измеренные значения сигнала в  $n = 0, \dots, N-1$ ):

$$F_{DFT} = \sum_{n=0}^{N-1} f[k] e^{-j \frac{2\pi}{N} nk} \quad (1)$$

2. Получившийся спектр преобразуется в мел-шкалу. Мел-шкала вводится для того, чтобы учитывать эффект деформации частот в соответствии с психоакустическими наблюдениями. Слуховая система человека проще различает разницу между более низкими частотами. Таким образом, высота звука, которая воспринимается человеком, нелинейно зависит от частоты. Ниже представлена одна из формул, которая выражает эту зависимость ( $f$  – частота (Гц)):

$$f_w = 1127.01048 \cdot \log \left( \frac{f}{700} + 1 \right) \quad (2)$$

3. Мел-частотные кепстральные (один из видов гомоморфной обработки сигналов) коэффициенты удовлетворяют выражению ( $\hat{O}_k$ ,  $k = 1, \dots, L$ , – это мощность на выходе фильтра  $k$ ):

$$c_m = \sum_{k=1}^L \left( \log(\hat{O}_k) \right) \cos \left[ m \left( k - \frac{1}{2} \right) \frac{\pi}{L} \right], m = 1, \dots, L. \quad (3)$$

В данной работе была использована полносвязная нейронная сеть, так как данные хранятся в одномерном массиве. Нейроны в полносвязном слое связаны со всеми нейронами в предыдущем слое, что означает, что каждый нейрон в полносвязном слое может взаимодействовать с любым нейроном в предыдущем слое.

Для скрытых слоев нейронные сети применяется нелинейная функция активации ReLU [2]. Она преобразует входное значение в диапазоне  $[0, \infty)$ . На финальном слое используется функция активации softmax [2]. Она принимает вектор действительных аргументов и преобразует их в вектор, компоненты которого заключены в диапазоне от 0 до 1, а их сумма равна 1. Во время обучения использовалась функция потерь перекрестная энтропия (cross entropy) [2]. Эта функция вычисляет разницу между исходным вектором и вектором, который был предсказан моделью. Перекрестная энтропия была использована для бинарной классификации.

Для реализации взаимодействия пользователя был написан веб интерфейс, позволяющий записать и отправить свою голосовую команду

для дальнейшего использования ее в модели нейронной сети. Для реализации данного клиента использовался язык JavaScript.

Функционал веб интерфейса достаточно мал, зато очень эффективен и понятен. Функционал ограничивается кнопкой, на которую пользователь нажимает и произносит свою голосовую команду.

Для принятия данных от клиента, был написан сервер на языке программирования Python с использованием библиотеки Flask. Основная задача сервера — принятие аудио данных от клиента и сохранение их. Для реализации сервера был использован декоратор route, который регистрирует URL — адрес, по которому будет доступно соответствующее представление, а также устанавливает правила маршрутизации входящих запросов.

Все вычисления были произведены на локальной машине, имеющий видеокарту GeForce GTX 1660 Ti с 6 гигабайтами памяти. Каждый подход был протестирован как минимум на 100 эпохах. Если в какой-то момент времени функция потерь начинала расти на протяжении нескольких эпох, обучение останавливалось. Далее подходы анализируются и сравниваются между собой.

Для сходимости простой модели, обученной с нуля, потребовалось около 400 эпох. Каждая из них занимала менее минуты, но при этом было задействовано достаточное количество времени и производительной мощности.

В конечном итоге, после 400 эпох, модель достигла 62% точности.

В таблице 1 представлены результаты обучения модели на разном количестве эпох.

**Таблица 1 — Результаты обучения модели**

Количество эпох	Процент точности
200	29%
300	32%
400	62%
1000	31%
2000	30%

В данном исследовании нам удалось получить точность предсказания 62%. Как мы можем видеть из таблицы 1, точность предсказания падает на большом количестве эпох, это связано с переобучением модели нейронной сети. Хочу отметить, что с увеличением скрытых слоев полносвязной нейронной сети, выросла точность предсказания в 1.6 раза.

Для улучшения модели и точности предсказания можно использовать другой метод предобработки, а именно STFT (оконное преобразование Фурье), и изменение на сверточную архитектуру нейронной сети.

Литература:

1. Mohammad Faisal, Aasim Ullah, Daud Shakil Automatic Isolated Speech Recognition System Using MFCC. - Saarbrücken: LAP LAMBERT Academic Publishing, 2014, pp. 34-38.
2. Andrew W. Trask Grokking Deep Learning. - United States of America: MANNING Shelter Island, January 2019. pp. 162-172, 258.

## **ПОВЫШЕНИЕ КАЧЕСТВА ИЗОБРАЖЕНИЯ С ПОМОЩЬЮ ГЛУБОКИХ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ**

**И.М. Попов**

**Новосибирский государственный технический университет,  
г. Новосибирск, i.porov.2018@stud.nstu.ru**

**Научный руководитель: Попов А. А., д.т.н., профессор**

*Основная цель данной работы – воспроизведение и сравнение глубоких нейронных сетей EDSR и WDSR, предназначенных для улучшения качества изображения. Реализованные нейронные сети обучались на наборе данных DIV2K и показали свою эффективность в обработке изображений.*

*The primary objective of this study is to replicate and compare deep neural networks, namely EDSR and WDSR, designed to enhance image quality. The implemented neural networks were trained on the DIV2K dataset, demonstrating their effectiveness in image processing.*

В современном обществе визуальное восприятие играет ключевую роль, и качество изображений становится неотъемлемой частью повседневной жизни и профессиональных областей. Нейронные сети для решения задачи супер-разрешения (Super-Resolution или SR) представляют собой перспективное направление, привлекая внимание исследователей и инженеров, и обещают решение проблем с качеством изображений, вызванных ограничениями оборудования и сжатием данных.

Мы сконцентрировались на воспроизведении таких архитектур нейронных сетей как EDSR [0] и WDSR [0], предназначенных для решения задачи супер-разрешения для одного изображения (Single Image Super-Resolution или SISR). Обе эти архитектуры являются глубокими остаточными сверточными нейронными сетями, в которых применяются остаточные блоки с пропускающими соединениями. Этот подход

позволяет сохранять признаки, которые могли бы быть потеряны в процессе свертки.

Концепция внедрения остаточных блоков с пропуском соединения уходит корнями в структуру SRResNet [0]. В отличие от данной архитектуры, EDSR отказывается от применения нормализации по батчам в остаточных блоках, в то время как WDSR воспользуется более широкими функциями активации и заменит нормализацию по батчам на нормализацию по весам. Эти отличия представляют собой важные архитектурные вариации, которые могут оказать влияние на эффективность и обобщающую способность нейронных сетей для задач супер-разрешения.

Для обучения реализованных нейронных сетей применялся набор данных DIV2K, включающий в себя обучающую выборку из 800 пар изображений высокого разрешения и их соответствующих сжатых вариантов. Для расширения набора производилось с помощью вырезки небольших фрагментов изображений, которые затем случайным образом поворачивались и отражались. Таким образом можно практически бесконечно генерировать различные обучающие данные. Оценка производительности моделей проводилась на тестовой выборке, включающей 100 пар изображений, также взятых из данного набора данных.

Нами были реализованы сети для четырехкратного увеличения изображения. Это верхний предел, который поддерживает архитектура EDSR, но WDSR способна на девятикратное увеличение. В таблице 1 приведены средние метрики пикового сигнала к фону (PSNR) индекса структурной схожести (SSIM) по тестовой выборке.

**Таблица 1 – Метрики для тестовой выборки**

	<b>PSNR</b>	<b>SSIM</b>
<b>EDSR</b>	28.890	0.8211
<b>WDSR</b>	28.906	0.8214

Как видно из таблицы в среднем WDSR работает немного эффективнее. Для большей наглядности приведем примеры работы для изображения с высокой (*Рис. 1*) и низкой степенью детализации (*Рис. 2*). На рисунках используются следующие обозначения: LR – исходное изображение низкого качества, ONLINE – работа онлайн сервиса, EDSR – работа EDSR, WDSR – работа WDSR. Для сравнения работы реализованных сетей был использован также онлайн сервис Vimdore [0]. Соответствующие метрики для этих изображений представлены в таблице 2.



Рис. 1 – Демонстрация фрагмента работы на высоко детализированном изображении.

**Таблица 2 – Метрики для единичных изображений**

	Высокая детализация (Рис. 1)		Низкая детализация (Рис. 2)	
	PSNR	SSIM	PSNR	SSIM
<b>EDSR</b>	28.07	0.8638	33.198	0.9096
<b>WDSR</b>	28.022	0.8641	33.212	0.9097
<b>ONLINE</b>	24.106	0.7411	29.076	0.839

Как мы можем видеть по метрикам, реализованные нами сети показывают результат более высокого качества, чем дают свободные источники в сети. Также заметим, что WDSR все так же показывает немного более высокое качество, чем EDSR.

Полученные результаты подчеркивают потенциал и эффективность применения глубоких остаточных сверточных нейронных сетей в контексте повышения качества изображений. Более того, выбранные нами архитектуры предоставляют перспективные варианты для дальнейших исследований в области супер-разрешения.

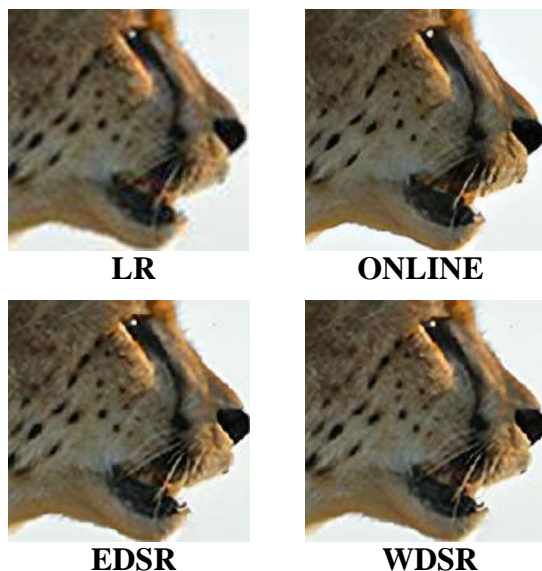


Рис. 2 – Демонстрация фрагмента работы на низко детализированном изображении

Литература:

1. Enhanced Deep Residual Networks for Single Image Super-Resolution / B. Lim et al. // arXiv, 2017. <https://arxiv.org/pdf/1707.02921.pdf>.
2. Wide Activation for Efficient and Accurate Image Super-Resolution / J. Yu et al. // arXiv, 2018. <https://arxiv.org/pdf/1808.08718.pdf>
3. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network / Ledig C. et al. // arXiv, 2017. <https://arxiv.org/pdf/1609.04802.pdf>
4. Vimdore : [сайт]. — URL: <https://www.vidmore.com/ru/image-upscaler/> (дата обращения: 10.11.2023).

## ПОРОЖДАЮЩИЕ СОСЯЗАТЕЛЬНЫЕ НЕЙРОННЫЕ СЕТИ ДЛЯ ДАННЫХ ВРЕМЕННЫХ РЯДОВ

**А.О. Хуснутдинов, В.С. Карманов**

**Новосибирский Государственный Технический Университет,  
г. Новосибирск, evolext@gmail.com**

**Научный руководитель: Хабаров В.И., д.т.н., профессор**

*В работе рассмотрена задача расширения малых наборов данных временных рядов с помощью генеративных моделей. Приведен краткий анализ существующих подходов и применение одного из методов к реальным данным.*

*The paper considers the problem of expanding small datasets of time-series using generative models. A brief analysis of existing approaches and application of one of the methods to real data are given.*

Анализ временных рядов имеет большую практическую значимость и может помочь определить, какие закономерности заключаются в наблюдаемых данных. Выявление принципов формирования значений временного ряда играет важное значение для составления моделей прогноза будущих значений временных рядов и принятия на их основе решений, таких как планирование производства, распределение ресурсов, план лечения пациента и другие приложения.

Точность прогнозных алгоритмов тесно связана с объемом доступных обучающих данных. Современные модели прогнозирования содержат десятки и сотни миллионов параметров. Следовательно, для надлежащего определения значений этих параметров требуется, как минимум, соответствующий объем обучающих данных. Однако для большинства приложений доступны лишь небольшие наборы размеченных данных, что часто приводит к низкой эффективности решаемой задачи. Получение большого объема значений временных рядов является актуальной проблемой, причины которой могут обуславливаться регулярными изменениями внешней среды, ограниченным набором вспомогательных признаков, высокой стоимостью сбора данных, длительными периодами в циклах изменения наблюдений.

Цель настоящей работы заключается в разработке способа генерации синтетических данных временных рядов на основе небольшого количества реальных наблюдений для возможности применения современных прогнозных моделей машинного обучения.

Дополнительный сбор данных не всегда является решением описанной проблемы, учет новых наблюдений может потребовать большого количества времени на поиск, сбор и подготовку. Другим

решением проблемы малых данных является аугментация, её суть заключается в создании разнообразных вариаций исходных данных путем внесения небольших изменений, чтобы обогатить обучающий набор и улучшить обобщающую способность модели. Чаще всего эта техника применяется в задачах компьютерного зрения, аугментация изображений может включать в себя вращение, изменение масштаба, отражения и другие преобразования [1]. Произвольные трансформации временных рядов могут исказить временные паттерны и внести случайные искажения, поэтому не применяются на практике.

С развитием нейронных сетей стали доступны более продвинутые решения с области генерации временных рядов, рекуррентные нейронные сети (Recurrent Neural Network, RNN) и их пошаговые предсказания использовались для генерации последовательностей в методах T-forcing [2] и P-forcing [3]. С появлением архитектуры порождающих состязательных сетей (Generative Adversarial Network, GAN) [4], были описаны различные методы их применения для генерации временных рядов. Метод C-RNNGAN [5] использует GAN с LSTM-сетями [6] в качестве генератора и дискриминатора, RCGAN [7] расширяет этот метод, усложняя структуру формирования выходов. Метод TimeGAN [8] является одним из современных подходов к генерации временных рядов, он использует четыре сетевые компоненты: embedding-сеть, recovery-сеть, дискриминатор и генератор.

Примем за  $X$  векторное пространство для временных признаков  $x_t$ , а  $H_x$  – векторное пространство скрытых состояний для временных признаков, тогда, embedding-сеть  $e: H_x \times X \rightarrow H_x$  переводит временные признаки в их скрытые состояния  $h_t = e(h_{t-1}, x_t)$ . В данном случае  $e$  вводится через рекуррентную сеть, а  $h_{t-1}$  – скрытое состояние для предыдущего элемента ряда. С другой стороны, recovery-сеть  $r: H_x \rightarrow X$  переводит элементы из пространства скрытого состояния в их восстановленные представления признаков  $x_t = r(h_t)$ , в данном случае  $r$  представляется сетью прямого распространения. Вместо прямого вывода синтетических данных в пространство признаков, сеть генератора выводит в пространство скрытых состояний  $Z_x$ . Сеть генерации  $g: H_x \times Z_x \rightarrow H_x$  переводит кортеж из случайных векторов и скрытых состояний в синтетические скрытые состояния  $h_t = g(h_{t-1}, z_t)$ , в данном случае  $g$  вводится через рекуррентную сеть, а  $z_t$  – случайный процесс. Наконец, сеть дискриминатора  $d: H_x \rightarrow [0, 1]$  получает



элементы пространства скрытых состояний и возвращает класс  $y_t$ , где обозначение  $y_*$  относится либо к реальным  $y_*$ , либо синтетическим  $y_*$  данным.

Сети `embedding` и `recovery` должны предоставлять точное восстановление  $x_t$  в исходные данные  $x_t$  из их скрытых состояний  $h_t$ , поэтому, первая целевая функция – ошибка восстановления:

$$L_R = E_{x \sim p} \left[ \left\| x_t - x_t \right\|_2 \right], \quad (1)$$

соответственно, неподконтрольная функция ошибки генератора и дискриминатора:

$$L_U = E_{x \sim p} \left[ \log(y_t) \right] + E_{x \sim p} \left[ \log(1 - y_t) \right] \quad (2)$$

где  $p$  – распределение исходных данных,  $p$  – распределение синтезированных данных.

Ориентир исключительно на бинарный отклик дискриминатора не обеспечит выявление условных распределений в данных, применим дополнительно принцип максимума правдоподобия [9] и получим подконтрольную функцию ошибки:

$$L_S = E_{x \sim p} \left[ \left\| h_t - g(h_{t-1}, z_t) \right\|_2 \right], \quad (3)$$

таким образом  $L_U$  стимулирует генератор создавать реалистичные последовательности,  $L_S$  дополнительно обеспечивает получение похожих переходов между моментами времени  $t$ .

В качестве исходных данных для генерации были выбраны ежедневные значения потребления тепловой энергии жилым домом за один календарный год [10] с 01.01.2011 по 01.01.2012. В выбранных данных отсутствует тренд, наблюдается, как минимум, один компонент сезонности. График значений представлен на *Рис. 1*.

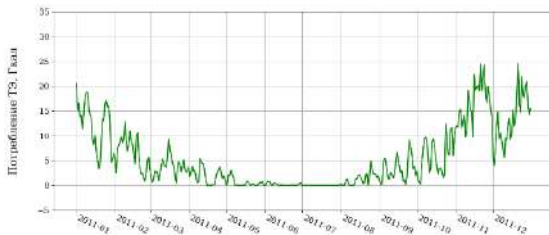


Рис. 3 – Объем потребления тепловой энергии за один год

На основе выбранных значений потребления сгенерируем моделью TimeGAN еще один набор из 365 значений и сравним его с исходным.

В отличие от генерации изображений, где достаточно визуальной оценки, для временных рядов необходимо анализировать диапазон принимаемых значений и сохранение временных зависимостей. График автокорреляционной функции (АКФ) позволяет нам оценить степень корреляции между разными значениями ряда, а гистограмма значений позволяет оценить распределение значение и установить, соответствует ли сгенерированный ряд исходному. Сравнение представлено на *Рис. 2*.

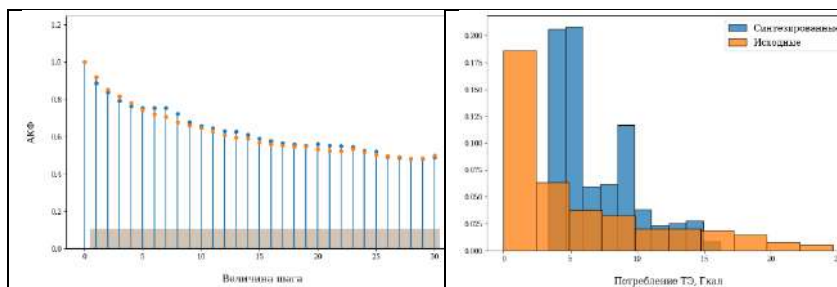


Рис. 2 – Визуальное сравнение сгенерированных значений ряда (синий цвет) с исходными (оранжевый цвет) с помощью (а) АКФ и (б) гистограммы

Вид АКФ говорит о сохранении зависимостей между значениями с сгенерированных данных, но несоответствие гистограмм указывает на то, что модель генерации плохо учитывает граничные значения ряда.

В работе изучалась задача генерации данных временных рядов при помощи порождающих состязательных сетей. В частности, приведены исследования модели TimeGAN, небольшой набор доступных методов и наличие несоответствий в результатах указывают на необходимость дальнейших исследований в данной области.

Литература:

1. A survey on Image Data Augmentation for Deep Learning. / C. Shorten, T. Khoshgoufar // *J Big Data*. – 2019. – Vol. 6, № 60. <https://doi.org/10.1186/s40537-019-0197-0>.
2. A. Graves Generating Sequences With Recurrent Neural Networks [Electronic resource] / A. Graves // arXiv [cs.NE] – Aug 2013. – 1308.0850v5. – Mode of access: <https://arxiv.org/abs/1308.0850>. – Title from screen.
3. Professor forcing: a new algorithm for training recurrent networks. / A. Goyal, A. Lamb, Y. Zhang, S. Zhang, A. C. Courville // *Advances in Neural Information Processing Systems*. – 2016. – № 29, pp. 4601-4609.
4. Generative adversarial networks. Advances in neural information processing systems. / I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley // *Advances in Neural Information Processing Systems*. – 2014. – № 27.
5. O. Mogren C-RNN-GAN: Continuous recurrent neural networks with adversarial training [Electronic resource] / O. Mogren // arXiv [cs.AI] – Nov 2016. – 1611.09904v1. – Mode of access: <https://arxiv.org/abs/1611.09904>. – Title from screen.
6. Long Short-term Memory / Sepp Hochreiter // *Neural Computation* – December 1997. – 9(8): 1735-1780 p. DOI: 10.1162/neco.1997.9.8.173.
7. C. Esteban Real-valued (Medical) Time Series Generation with Recurrent Conditional GANs [Electronic resource] / C. Esteban, S.L. Hyland, G. Rätsch // arXiv [stat.ML] – Jun 2017. – 1706.02633v2. – Mode of access: <https://arxiv.org/abs/1706.02633>. – Title from screen.
8. Time-series generative adversarial networks. / J. Yoon, D. Jarrett, M. Schaar // *Advances in neural information processing systems*. – 2019. – № 32.
9. Постовалов С. Н. Математическая статистика. Конспект лекций : учеб. пособие / С.Н. Постовалов, Е.В. Чимитова, В.С. Карманов ; – 2-е изд. – Новосибирск: Изд-во НГТУ, 2017. – 140 с.
10. 8 years of hourly heat and electricity demand for a residential building [Electronic resource] – Mode of access: <https://ieee-dataport.org/open-access/8-years-hourly-heat-and-electricity-demand-residential-building> (date of the application: 01.10.2023). – Title from screen.

# ГЕНЕРАЦИЯ ОБУЧАЮЩИХ ПРИМЕРОВ С ИСПОЛЬЗОВАНИЕМ СОСТЯЗАТЕЛЬНЫХ НЕЙРОННЫХ СЕТЕЙ

**Е.С. Шеховцов**

**Новосибирский государственный технический университет,  
г. Новосибирск, shexovczov.2018@stud.nstu.ru**

**Научный руководитель: Попов А.А., д.т.н., профессор**

*Цель работы – создание нейронной сети глубокого обучения, выполняющей задачу моделирования синтетических данных для обучения других моделей машинного обучения. В качестве основы использовались две архитектуры генеративно-сопоставительных моделей – DCGAN и SAGAN. Для взятых моделей были проведены эксперименты по выбору типа начального шума и сравнение с классическими методами аугментации изображений. Для проверки использовалась задача классификации, а качество полученных синтетических данных определялось по результатам работы классификатора. Полученные сети были протестированы, а результаты проанализированы на двух различных наборах данных: с фотографиями человеческих лиц и с рентгенограммами органов грудной клетки с различными видами заболеваний.*

*The aim of the article is to create a deep learning neural network that performs the task of modelling synthetic data for training other machine learning models. Two architectures of generative-adversarial models, DCGAN and SAGAN, were used as a basis. For the models taken, experiments were conducted to select the type of initial noise and compared with classical image augmentation methods. A classification task was used for validation and the quality of the obtained synthetic data was determined by the results of the classifier. The resulting networks were tested and the results were analyzed on two different datasets: with photographs of human faces and with chest radiographs with different types of diseases.*

Синтетические данные – это информация, которая создается на компьютере для дополнения или замены реальных данных с целью улучшения моделей искусственного интеллекта (ИИ), защиты конфиденциальных данных и уменьшения предвзятости нейросетей.

Проблема недостатка данных для обучения моделей машинного обучения остается актуальной. Одним из наиболее часто встречаемых методов обучения нейронных сетей является обучение с учителем, а оно требует большое количество обучающих примеров. При этом далеко не для каждой задачи можно найти подходящие объемы информации. Например, медицинская диагностика имеет либо недостаточный объем данных, либо личную информацию, которую нельзя распространять. Для решения таких задач существуют различные методы для создания синтетических данных, наиболее популярные из которых –

классические алгоритмы аугментации данных и генеративные модели [1][2].

Целью данной работы являлось написание программы, реализующей генерацию большого количества обучающих примеров для выборки оригинальных данных малого размера. Генеративные модели достигли невероятного качества, но оно доступно только при значительных вычислительных мощностях. В качестве основы для будущих нейронных сетей были взяты две модели – DCGAN и SAGAN [3]. Первая модель является одной из самых популярных по частоте использования, а последняя – имеет в себе механизм внимания (англ. attention), который призван обратить внимание генеративной модели на важные части исходных данных.

Созданная модель на основе архитектуры DCGAN является прямым расширением GAN, за исключением того, что она использует сверточный и сверточно-транспонирующий слои в дискриминаторе и генераторе соответственно. Модель на основе архитектуры SAGAN использует слой трансформер, выполненный в форме нелокальной нейронной сети и использованием механизма самовнимания (англ. self-attention). Дискриминаторы обеих сетей используют функцию активации LeakyReLU.

Для оценки качества синтетических данных используется дополнительная нейросетевая модель глубокого обучения для классификации. Модель представляет собой предобученную модель EfficientNet, которая находится в открытом доступе. При этом последний уровень нейронной сети – полносвязный слой – был изменен согласно количеству классов для исходных наборов данных, а остальные слои были заморожены. Для оценки качества модели в данной работе использовались следующие метрики: ошибка и F-мера.

Наборы данных, которые использовались для проверки и тестирования – набор изображений лиц Celeba [4] и набор рентгенограмм внутренних органов груди Lung Diseases X-Rays [5].

Сравнение происходило для разных соотношений реальных данных, синтетических данных, созданных с помощью генеративно-состязательной нейронной сетью и синтетических данных, полученных с помощью стандартной данных аугментации. Результаты сравнения представлены на рисунке 1.

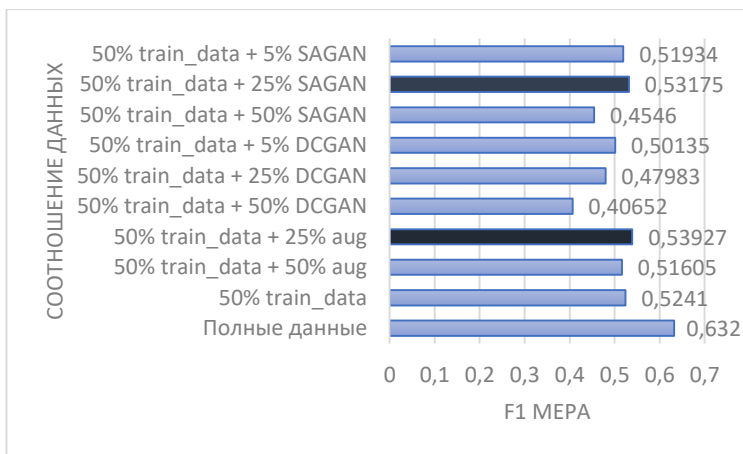


Рис. 4 – Оценка качества классификации для разных соотношений реальных и синтетических данных

Для реализованных моделей дополнительно был поставлен эксперимент по выявлению наиболее подходящего типа шума, который подается на вход генератора для создания новых изображений. В сравнении участвовали: шум Гаусса, шум Бернулли, шум соли и перца.

Результаты сравнения показаны в таблице 1.

Таблица 3 – Сравнение качества классификации при использовании разных видов шума

Архитектура	Модель случайного шума	Метрика	
		F-мера	Ошибка
DCGAN	Гауссовский шум	0.47983	1.9983
	Шум Бернулли	0.39284	2.3852
	Шум соли и перца	0.37284	2.4294
SAGAN	Гауссовский шум	0.53175	1.5795
	Шум Бернулли	0.45042	2.0142
	Шум соли и перца	0.42042	2.1142

Для улучшения результатов работы в дальнейшем можно обучать сети на значительно более мощном оборудовании с большими размерами батчей и увеличенным числом слоев. Изменение архитектуры также может повлиять на результаты работы и качестве выходных изображений.

Также выдвигается гипотеза, что успех в генерации синтетических данных может зависеть от того, насколько они повышают репрезентативность обучающего набора данных. Если синтетические

данные «повторяют» основные свойства в среднем, то эффекта от их использования может не быть. Для проверки этой гипотезы в дальнейшем предлагается оценивать репрезентативность исходного набора данных и набора с добавлением синтетических данных. Дополнительные эксперименты следует провести на разных наборах данных из разных областей, чтобы сделать наиболее полный вывод на это предположение.

Полученные результаты показывают, что использование генеративно-состязательных сетей имеет потенциал, но требует точного подбора параметров. Архитектура SAGAN со слоями внимания позволили достичь точности классификации почти сравнимым с классическими методами аугментации.

Сравнение разных типов шумов показало, что использование шума по умолчанию – Гауссовского шума является наиболее эффективным.

В конечном итоге можем сказать, что поставленные в работе задачи частично выполнены и работа будет продолжаться. Написанная программа была протестирована и проанализирована. В том числе были предложены варианты улучшения качества получаемых результатов. Реализованная нейронная сеть, а также полученные выводы позволяют сделать шаг в сторону дальнейшего развития методов эффективного создания синтетических данных.

#### Литература:

1. Goodfellow I.J. Generative Adversarial Nets // NIPS. - 2014. - P. 2672–2680.
2. Mao X. Least Squares Generative Adversarial Networks // ICCV. IEEE Computer Society. - 2017. - P. 2813–2821.
3. Radford A., Metz L., Chintala S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks // ICLR. – 2016.
4. Large-scale CelebFaces Attributes (CelebA) Dataset: [Electronic resource]. URL: <http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html> (дата обращения: 17.05.22).
5. Lung Diseases X-Rays Dataset: [Electronic resource]. URL: <https://qims.amegroups.org/article/view/5132/html> (дата обращения: 13.12.22).

**Секция ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ,  
СТАТИСТИКА И ЭКОНОМЕТРИКА**

**ИССЛЕДОВАНИЕ ФАКТОРОВ ЗАВИСИМОСТИ ДАННЫХ ЭЭГ  
И МАТЕМАТИЧЕСКИХ СПОСОБНОСТЕЙ ИСПЫТУЕМЫХ**

**С.С. Альков**

**Новосибирский государственный технический университет,**

**г. Новосибирск, Nemul2017@yandex.ru**

**Научный руководитель: Тимофеева А.Ю., к.э.н**

*В данной работе был проведен анализ данных ЭЭГ с целью нахождения факторов, оказывающих влияние на математические способности. По результатам исследования была построена модель бинарной классификации, способная определить уровень математических способностей.*

*In this paper, EEG data was analyzed in order to find factors that influence mathematical ability. Based on the results of the study, a binary classification model capable of determining the level of mathematical ability was constructed.*

Данные, использованные в настоящем исследовании, были взяты из более крупного исследования, направленного на сбор данных об активности мозга участников, выполняющих задания, связанные с когнитивной нагрузкой.

Всего в исследовании первоначально приняли участие 36 не имеющих проблем со здоровьем добровольцев (27 женщин и 9 мужчин). Все участники — студенты 1-3 курсов Киевского национального университета имени Тараса Шевченко (Учебно-научный центр "Институт биологии и медицины" и факультет психологии) в возрасте от 18 до 26 лет (средний возраст = 18.25 лет, стандартное отклонение (SD) = 2.14 лет). Участники могли участвовать в исследовании, если у них была нормальная или с коррекцией до нормы острота зрения, нормальное восприятие цветов и не было клинических проявлений психических или когнитивных расстройств, вербальных или невербальных нарушений обучаемости. Критериями исключения были употребление психоактивных препаратов, наркотическая или алкогольная зависимость, а также психиатрические или неврологические жалобы [1]. Полные статистические данные об испытуемых приведены в таблице 1.

В данном исследовании для определения математических способностей использовались арифметические примеры, заключающиеся в вычитании одного числа (двухзначного) из другого (четырёхзначного). При этом числа передавались испытуемым путем



устной коммуникации. В зависимости от числа решенных примеров, участникам присваивались две категории (0 — меньше 10 решенных примеров, 1 — больше 10).

**Таблица 1 — Данные об участниках**

Количество участников	36
Процент мужчин	25%
Процент женщин	75%
Средний возраст	18.25
Стандартное отклонение по возрасту	2.14

Данные ЭЭГ были собраны для изучения определенных частотных диапазонов (альфа, бета и гамма). Они были записаны при помощи 21-канальных электродов, распределенных по коже головы по схеме 10/20. Данные сэмпировались с частотой 500 Гц.

Для удаления шума, вызванного электроникой из окружающей среды, использовался фильтр с полосой пропускания 50 Гц. Для устранения артефактов был использован метод независимых компонентов [1].

Частотные диапазоны, используемые в текущем исследовании, определены следующим образом: альфа (8,0-12,0 Гц), бета (12,0-30,0 Гц), гамма (>30,0 Гц)

Изначальные ЭЭГ-данные хранились в edf формате, для их чтения был использован mne, пакет с открытым исходным кодом для изучения, визуализации и анализа нейрофизиологических данных человека для языка Python (в исследовании использовалась версия 3.9). Для визуализации и графиков ЭЭГ-данных и метрик, связанных с ними, использовался модуль mne.viz.

Для каждого участника исследования имелось 2 файла с данными — один был записан непосредственно во время математических вычислений, другой — в состоянии покоя. Обе записи имеют одинаковую длину в 60 секунд. Для построения модели использовались данные, записанные во время вычислений. Они были разбиты на равномерные эпохи длиной в 2 секунды.

В целях проведения расчета метрик коннективности был использован пакет mne\_connectivity, предназначенный для вычисления коннективности и связанных с ней показателей данных МЭГ, ЭЭГ или ИЭЭГ. Из данного пакета были использованы 5 видов метрик коннективности — когерентность (coh), мнимая когерентность (imcoh), индекс фазового лага (pli), направленный индекс фазового лага (dpli), взвешенный индекс фазового лага (wpli).

В итоге было получено 15 датасетов (по одному для каждого из трех спектров частот для каждой из пяти видов метрик коннективности), каждый размером 36x210 (ввиду симметричности метрик), таким образом, количество параметров сильно превышает количество имеющихся наблюдений.

Далее по полученным матрицам с помощью библиотеки scikit-learn для языка Python была проведена логистическая регрессия с L2-регуляризацией. Целью данной регрессии являлось предсказание математических способностей испытуемого с помощью значений метрик коннективности между определенными каналами записи ЭЭГ-данных, а следовательно, и между соответствующими областями головного мозга [3].

Для выбора лучшей модели с минимально необходимым набором факторов использовалась 5-блочная кросс-валидация, призванная не допустить переобучение путем усреднения результатов модели по всем имеющимся данным [2].

В результате исследования всех комбинаций различных видов метрик коннективности и диапазонов альфа, бета и гамма частот, для построения модели были выбраны факторы, обеспечивающие наибольшее значение  $R^2$ . Таковыми факторами оказались:

— матрица когерентности, построенная на частотах гамма диапазона ( $R^2 = 0.18$ ). На рисунке 1 приведены значения коэффициентов данной модели;

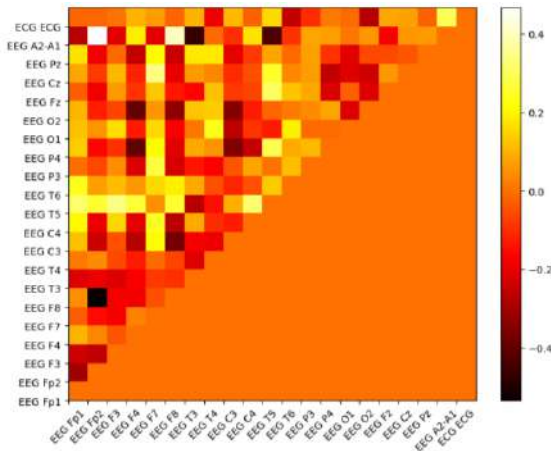


Рис.1 — Значения коэффициентов модели для матрицы когерентности, построенной на частотах гамма диапазона



Далее полученные факторы были собраны в одну модель логистической регрессии, качество которой было оценено по метрике F1 score, а также по точности.

F1 score: 0.8229

Точность: 0.725

Таким образом, отобранные факторы позволяют проводить по ним весьма надежную классификацию математических способностей испытуемых даже с использованием простой линейной модели

Литература:

1. Zyma I., Tukaev S., Seleznov I. Electroencephalograms during Mental Arithmetic Task Performance // MDPI. — 2018. — Vol. 4, N 1. — P. 1-6.

2. Nuamah J. K., Seong Y. Support vector machine (SVM) classification of cognitive tasks based on electroencephalography (EEG) engagement index // Brain-Computer Interfaces. — 2017. — Vol. 5, N 1. — P. 1-12.

3. Timofeeva A. Y., Murtazina M. S. Feature selection for EEG data based on logistic regression // Actual problems of electronic instrument engineering (APEIE–2021) : proc. of the 15 intern. sci. and techn. conf., Novosibirsk, 19–21 Nov. 2021. – Novosibirsk : Publ. NSTU, 2021. – P. 604–609. – ISBN 978-1-6654-3408-9.

## **ИССЛЕДОВАНИЕ МЕТОДОВ ИЗМЕРЕНИЯ ФРАКТАЛЬНОЙ РАЗМЕРНОСТИ ДЛЯ СЕГМЕНТАЦИИ ПЕРЕХОДНОГО СЛОЯ ПРИ SPS-СПЕКАНИИ**

**Н.Е. Бортник**

**Новосибирский государственный технический университет,  
г. Новосибирск, bortnikovnik@gmail.com**

**Научный руководитель: Тимофеев В.С., д.т.н., доцент**

*В статье рассматриваются два метода измерения фрактальной размерности на изображении: метод кубов и метод треугольных призм. По каждому методу построены поля фрактальных размерностей для снимка сплава двух смесей металлов. Методы сравниваются по визуализации построенных полей фрактальных размерностей и по параметрам распределений, которые характеризуют эти поля.*

*The article deals two methods of measuring fractal dimension in an image: the cube method and the triangular prisms method. For each method, fields of fractal dimensions were formed for the image of an alloy of two mixtures of metals. The methods are compared by visualizing the formed fields of fractal dimensions and by the parameters of distributions that characterize these fields.*

## **Введение**

Исследователи в области материаловедения занимаются созданием и изучением различных материалов. Один из способов получения нового продукта — это SPS-спекание двух смесей металлов в порошковом виде. В процессе такого спекания обе смеси меняют свою структуру, а между ними образуется переходный слой. Ширина переходного слоя влияет на свойства конечного продукта: с увеличением ширины переходного слоя увеличивается хрупкость материала. Поэтому задача измерения ширины переходного слоя имеет серьёзное значение в этих исследованиях.

Один из способов исследований конечного продукта – изучение сплава с помощью снимков его микроструктуры. На таких снимках можно отчётливо увидеть переходный слой, так как обе смеси металлов и переходный слой отличаются своей текстурой. В данный момент исследователи измеряют ширину переходного слоя ручным методом, что занимает достаточно много времени и не гарантирует высокой точности измерения. Этот процесс можно автоматизировать для ускорения исследований.

Первым шагом в автоматическом измерении ширины переходного слоя является его сегментация на снимках сплавов. Для решения задачи текстурной сегментации был выбран фрактальный подход, так как он хорошо работает с естественными текстурами [1].

В рамках данной работы была поставлена цель сравнить методы измерения фрактальных размерностей на изображении и выбрать метод, который наиболее отчётливо определяет области каждой текстуры на снимке сплава.

### **Постановка задачи**

Снимки сплавов двух смесей металлов являются чёрно-белыми изображениями. На рисунке 1 пример такого изображения. Для анализа изображения используется оконная обработка, одним из параметров которой является ширина измерительного окна  $w$ . С её помощью строится поле фрактальных размерностей (ПФР) - матрица, размеры которой эквивалентны размерам изображения и элементы являются значением фрактальной размерности в соответствующем участке изображения [2].

Задача заключается в том, чтобы выбрать метод измерения фрактальной размерности на изображении, который наименее подвержен искажению оценки размерности, и с использованием которого компоненты в смеси нормальных распределений, характеризующее ПФР, наиболее различимы: больше разность между математическими ожиданиями.

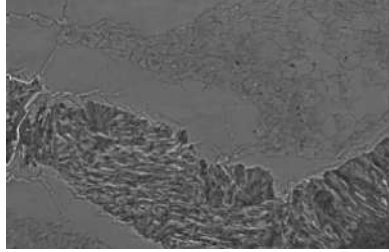


Рис. 1 – Пример снимка сплава двух смесей металлов.

### Методы измерения фрактальной размерности

При оценке фрактальной размерности в некоторой точке изображения строится линейная регрессия [3]:

$$\ln f(L) = \theta_0 + \theta_1 \ln \delta,$$

где  $f(L)$  – некоторая характеристика участка изображения  $L$  размера  $w$ , зависящая от метода измерения фрактальной размерности,  $\delta$  – размер субокна. Субокно – некоторая часть измерительного окна. Его размеры могут принимать значения, на которые  $w$  делится без остатка.

Фрактальная размерность оценивается по формуле:

$$D = f(\hat{\theta}_1),$$

где определение  $f(\hat{\theta}_1)$  зависит от метода измерения фрактальной размерности.

Рассмотрим два метода измерения фрактальных размерностей: метод кубов и метод треугольных призм.

В методе кубов:  $f(L)$  – количество кубов с ребром  $\delta$ , необходимых для покрытия поверхности участка изображения  $L$ ;  $f(\hat{\theta}_1) = \hat{\theta}_1$  [2].

В методе треугольных призм:  $f(L)$  – суммарная площадь верхних граней треугольных призм, построенных в каждом субокне;  $f(\hat{\theta}_1) = 2 - \hat{\theta}_1$  [4].

### Результаты построения полей фрактальных размерностей

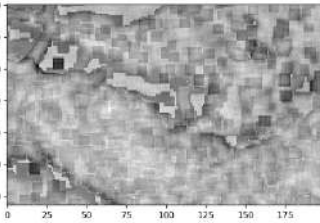
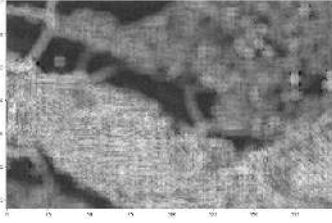
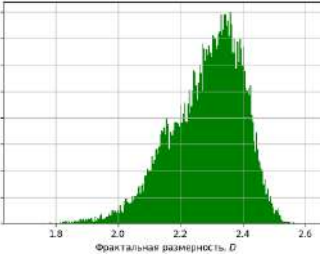
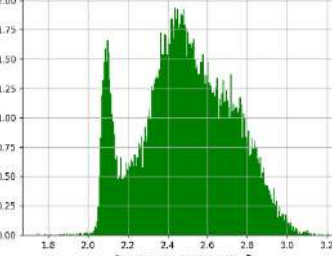
С помощью каждого метода были построены ПФР для снимка сплава на рисунке 1. Ширина измерительного окна равна 32 пикселя. Оценки параметров компонент в смеси распределений были получены с помощью EM-алгоритма [5]. Так как снимок имеет три уникальные текстуры, смесь делилась на три компоненты. Результаты применения обоих методов представлены в таблице 1.

При сравнении визуализации ПФР отчётливо видно, что ПФР, построенное по методу кубов, имеет достаточно много областей, имеющих прямоугольную форму, которых нет на исходном изображении и на визуализации ПФР по методу треугольных призм.

Гистограмма распределения размерностей, полученных по методу кубов, имеет два пика, в то время как при методе треугольных призм – 3.

Математические ожидания смежных компонент в смеси при методе кубов очень близки к друг другу: их средняя абсолютная разность равна примерно 0.12. При методе призм средняя абсолютная разность составила примерно 0.31.

**Таблица 1 – Сравнение ПФР, построенных по двум методам**

	Метод кубов	Метод треугольных призм
Визуализация ПФР		
Распределение размерностей		
Оценки параметров компонент в смеси	$\hat{\mu}_1 = 2.1382$ $\hat{\sigma}_1^2 = 0.0930$ $\hat{\mu}_2 = 2.2663$ $\hat{\sigma}_2^2 = 0.0608$ $\hat{\mu}_3 = 2.3784$ $\hat{\sigma}_3^2 = 0.0545$	$\hat{\mu}_1 = 2.1032$ $\hat{\sigma}_1^2 = 0.0308$ $\hat{\mu}_2 = 2.4231$ $\hat{\sigma}_2^2 = 0.1333$ $\hat{\mu}_3 = 2.7246$ $\hat{\sigma}_3^2 = 0.1327$

### Заключение

В данной работе были рассмотрены два метода измерения фрактальной размерности на чёрно-белом изображении: метод кубов и метод треугольных призм. С помощью каждого метода были построены поля фрактальных размерностей для снимка сплава двух смесей металлов. При сравнении визуализации ПФР на наличие искажений и сравнении абсолютных разностей математических ожиданий компонент в смеси распределений можно сделать вывод, что для дальнейшего

анализа снимков сплавов двух смесей металлов, полученных после SPS-спекания, лучше использовать метод треугольных призм.

Литература:

1. Анализ методов сегментации текстурных областей изображений в системах обработки изображений / Шитова О.В., Пухляк А.Н., Дроб Е.М. // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. - 2014. - №8 (179). - С. 182-188.
2. Тренихин В. А. Методы фрактальной обработки и комплексирования радиолокационных и спектрально-аналитических данных в системах космического наблюдения: дис. канд. т. н. наук: 05.12.14. - Екатеринбург, 2017. - 123 с.
3. Фисенко В. Т., Фисенко Т. Ю. Фрактальные методы сегментации текстурных изображений // Известия высших учебных заведений. Приборостроение. - 2013. - №5. - С. 63-70.
4. Fractal analysis of remotely sensed images: A review of methods and applications / W. Sun, G. Xu, P. Gong, S. Liang // International Journal of Remote Sensing. - 2006. - N 27. - P. 4963-4990.
5. Nguyen, L. Tutorial on EM Algorithm. Preprints 2018, 2018020131. - doi: 10.20944/preprints201802.0131.v9

## **ИССЛЕДОВАНИЕ ИНДИКАТОРОВ ВОСТРЕБОВАННОСТИ ПРОФЕССИЙ В НОВОСИБИРСКОЙ ОБЛАСТИ**

**В.К. Дворецкая**

**Новосибирский государственный технический университет,**

**г. Новосибирск, dvoreczka.2018@stud.nstu.ru**

**Научный руководитель: Тимофеев В.С., д.т.н., доцент**

*Проведено исследование динамики кадровой потребности по Новосибирской области в период с 2020 по 2022 год. Построены графики количества вакансий и средней заработной платы по Новосибирской области. Результаты исследования планируется внедрить в разработанную информационную систему для последующего моделирования.*

*A study of the dynamics of staffing needs for the Novosibirsk region in the period from 2020 to 2022 was conducted. Charts of the number of vacancies and average wages in Novosibirsk region have been constructed. It is planned to implement the results of the study into the developed information system for further modelling.*



С целью выравнивания спроса и предложения на рынке труда в Новосибирской области была разработана государственная информационная система «Платформа прогнозирования региональной кадровой потребности и планирования объема подготовки кадров по программам среднего профессионального образования и высшего образования» (далее - ГИС КП НСО) [1]. В рамках данной системы было проведено исследование, результаты которого планируются использоваться в построении моделей для ГИС КП НСО.

В отслеживании динамики изменения кадровой потребности особый интерес представляют сайты с информацией о вакансиях, которые являются косвенным индикатором востребованности профессии [2-4]. Основной информацией о вакансиях, которые были необходимы для исследования, послужили количество вакансий и средняя заработная плата по разделам ОКВЭД за 2020–2022 год [5].

В результате проведения агрегации данных по Новосибирской области были получены такие данные:

- количество открытых и закрытых вакансий по разделам ОКВЭД за период с 2020 по 2022 год;
- средняя заработная плата открытых и закрытых вакансий по разделам ОКВЭД за период с 2020 по 2022 год.

По результатам исследования была получена динамика количества вакансий по разделам ОКВЭД. Можно выделить два вида тенденции и три основных типа динамик, которые чаще всего встречаются при построении графиков изменения количества вакансий в течение двух лет.

На рисунке 1 приведены основные типы динамик и выделен один раздел ОКВЭД, как представитель типа динамики.

Большинство динамик количества вакансий относятся к возрастающей тенденции, на рисунке 1 приведены такие ее типы как:

- параболический тип, его представителем является раздел А – «Сельское, лесное хозяйство, охота, рыболовство и рыбоводство».
- логарифмический тип, его представителем является раздел F – «Строительство».
- линейный тип, его представителем на рисунке 1 является раздел N – «Деятельность административная и сопутствующие дополнительные услуги».

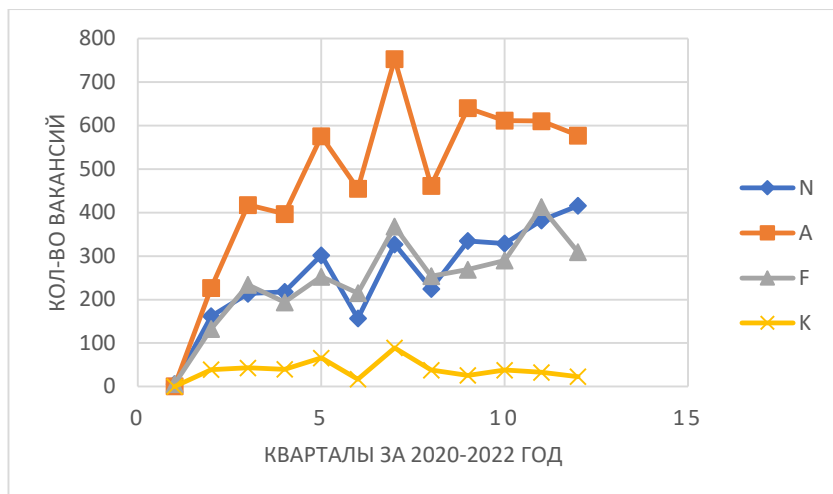


Рис. 1 – Динамика количества вакансий по разделам ОКВЭД

Основным представителем нисходящей тенденции на рисунке 1 является раздел К – «Деятельность финансовая и страховая».

У приведенных на рисунке 1 графиков динамики количества вакансий разделов N, F и K колебания первой части ряда незначительны, а тенденция характеризуется увеличением темпов развития. В свою очередь раздел А в начале ряда характеризуется резким темпом роста количества вакансий. Середина рядов N, А и F характеризуется быстрым увеличением темпов прироста и значительным уровнем колебаний. График раздела К в середине имеет значительные уровни колебаний, и общая тенденция характеризуется уменьшением количества вакансий. В третьей части рядов А, F и К тенденция характеризуется снижением темпов развития и уровни колебаний незначительны. А раздел N характеризуется увеличением темпов развития и уровни колебаний незначительны.

Для определения динамики востребованности профессий, необходимо рассмотреть изменение средней заработной платы с течением времени. Данный индикатор позволяет отметить рост заработных плат в определённых профессиях, что может указывать на их растущую важность и спрос на рынке труда.

На рисунке 2 приведены графики разделов ОКВЭД, которые отображают изменение средней заработной платы разделов ОКВЭД по кварталам 2020–2022 года.

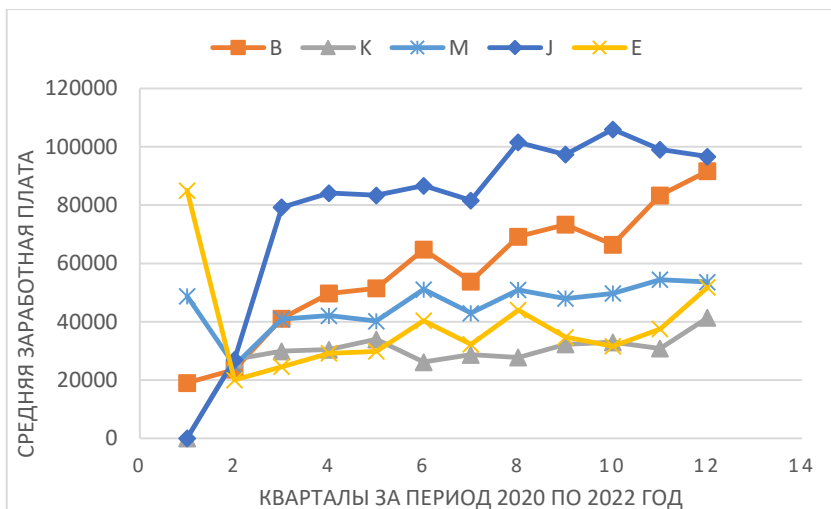


Рис. 2 – Динамика средней заработной платы по разделам ОКВЭД

Рассматриваемые разделы ОКВЭД: В – «Добыча полезных ископаемых», М – «Деятельность профессиональная, научная и техническая», Е – «Водоснабжение; водоотведение, организация сбора и утилизации отходов, деятельность по ликвидации загрязнений», Ж - «Деятельность в области информации и связи», К - «Деятельность финансовая и страховая».

После исследования средней заработной платы по разделам ОКВЭД за период с 2020 по 2022 год, были получены следующие результаты:

- наибольшую среднюю заработную плату имеет раздел Ж - 78 548,18 руб.
- наименьшую среднюю заработную плату имеет раздел К - 28 430,4 руб.

Проанализировав построенные графики на рисунке 2, можно сделать выводы, что в первой части ряда разделов Е и М есть резкое снижение средней заработной платы, во второй части ряда графики имеют увеличение средней заработной платы и уровни колебаний незначительные. Середина ряда разделов Е и М характеризуются возрастающей тенденцией и значительными колебаниями. В последней части ряда раздела Е после резкого снижения темпов развития с середины 2022 года (10 квартал) ряд имеет увеличение средней заработной платы. В свою очередь последняя часть ряда раздела М характеризуется линейным ростом заработной платы.

В графиках средней заработной платы разделов В и К наблюдается линейный рост заработной платы и незначительные колебания в

квартальных интервалах с начала 2020 по конец 2022 года. График изменения средней заработной платы раздела J в первой части имеет резкое увеличение средней заработной платы. Во второй части ряда тенденция остается возрастающей и уровни колебаний незначительны. В последней части ряда наблюдается сначала резкое увеличение средней заработной платы до середины 2022 года (10 квартал), а затем наблюдается резкое снижение средней заработной платы и колебания незначительны.

В дальнейшем планируется результаты исследования внедрить в информационную систему КП НСО для последующего построения математических моделей.

Литература:

1. Платформа прогнозирования региональной кадровой потребности и планирования объема подготовки кадров по программам среднего профессионального образования и высшего образования // ГИС КП НСО URL: <https://kr.nso.ru> (дата обращения: 07.11.2023).
2. Работа в Новосибирске, поиск персонала и публикация вакансий // HeadHunter URL: <https://hh.ru> (дата обращения: 07.11.2023).
3. Работа в России, свежие вакансии в России // SuperJob URL: <https://russia.superjob.ru> (дата обращения: 07.11.2023).
4. Работа России Общероссийская база вакансий и резюме // ТрудВсем URL: <https://trudvsem.ru> (дата обращения: 07.11.2023).
5. Расшифровка кодов ОКВЭД и их классификация 2023 год // Коды ОКВЭД URL: <https://код-оквэд.рф> (дата обращения: 29.10.2023).

## **ПРИМЕНЕНИЕ МЕТОДОВ НЕЧЕТКОЙ ЛОГИКИ ДЛЯ АНАЛИЗА ПРЕСТУПНОГО ПОВЕДЕНИЯ ПО ДАННЫМ ЭЭГ**

**И.А. Кутузова**

**Новосибирский государственный технический университет,  
г. Новосибирск, [kutuzova.2018@stud.nstu.ru](mailto:kutuzova.2018@stud.nstu.ru)  
Научный руководитель: Авдеенко Т.В., д.т.н., профессор**

*В настоящей работе получены результаты применения методов нечеткой логики для решения задачи классификации людей по данным ЭЭГ. Классифицирующий признак в исследовании может принимать два лингвистических значения: «преступники» и «несудимые». Выявлены зоны головного мозга и частотные диапазоны, в наибольшей степени определяющие исследуемое разбиение, а также, благодаря нечеткому разбиению, получены*

*вероятности отнесения испытуемого к определенному лингвистическому значению.*

*In this work, we obtained the results of using fuzzy logic methods to solve the problem of classifying people based on EEG data. The classifying feature in the study can take on two linguistic meanings: "criminals" and "unconvicted." The areas of the brain and frequency ranges that most determine the partition under study have been identified, and, thanks to the fuzzy partition, the probabilities of assigning a subject to a certain linguistic meaning have been obtained.*

Изучение темы проводилось по лекциям и статьям о методах нечеткой логики и создании нечетких деревьев принятия решений.

Целью данной работы является поиск отличительных особенностей у двух видов испытуемых (преступников и несудимых граждан) с помощью классификации их данных ЭЭГ методом нечеткого дерева решений.

Анализируемая выборка содержит записи двух категорий испытуемых: несудимые (норма) и преступники (заклученные с агрессивным и неагрессивным поведением).

В работе были использованы следующие показатели ЭЭГ:

- Аср – средняя амплитуда спектра в частотном диапазоне;
- Аmax – максимальная амплитуда спектра в частотном диапазоне;
- Аогиб - огибающая амплитуды спектра в частотном диапазоне;
- Fср - средневзвешенная частота;
- Fmax - частота максимальной по амплитуде гармоники.

Проведенное в прошлой работе исследование показало хорошие результаты в выявлении различий между классами у преступников в целом и несудимых людей. Было показано, методы классификации «дерево решений» и «случайный лес» могут быть применены в исследованиях на тему классификации агрессивного и преступного поведения людей на основе анализа данных ЭЭГ [1].

В данной работе исследование проводилось при помощи методов нечеткой логики [2-4]. Был реализован механизм построения нечеткого дерева решений с глубиной в 1 или 2. В качестве функции принадлежности была выбрана трапецевидная. Данные могли принимать три значения: "низкий", "средний" и "высокий". На Рисунке 2 представлен пример степеней принадлежности примеров к значениям атрибутов. Общий вид функции для атрибутов показан на Рисунке 1.

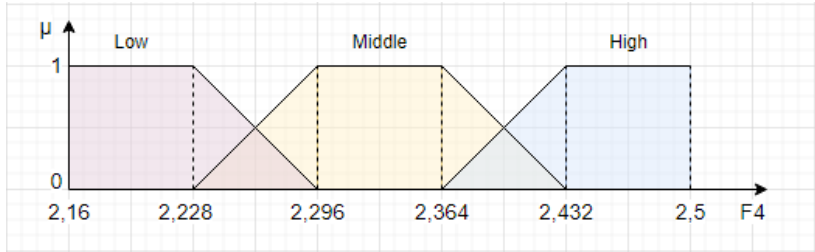


Рис. 1 – Пример трапецевидной функции принадлежности для атрибута F4.

Формула трапецевидной функции принадлежности ( $a, b, c, d$  — некоторые числовые параметры, принимающие произвольные действительные значения и упорядоченные отношением:  $a < b < c < d$ ; для Рисунка 1 и значения «средний»:  $a = 2.228, b = 2.296, c = 2.364, d = 2.432$ , а  $x$  — некоторое значение, которое может принимать данный атрибут):

$$\mu(x, a, b, c, d) = \begin{cases} 0, & \text{если } x \leq a; \\ \frac{x - a}{c - a}, & \text{если } a < x < c; \\ 1, & \text{если } c \leq x \leq b; \\ \frac{b - x}{b - d}, & \text{если } d < x < b; \\ 0, & \text{если } x \geq b; \end{cases} \quad (1)$$

	Fp1-L	Fp1-M	Fp1-H	Fp2-L	Fp2-M	Fp2-H
1	0.0000000	1.000000000	0.00000000	0.0000000	1.00000000	0.0000000
2	0.0000000	0.785942492	0.21405751	0.0000000	0.00000000	1.0000000
3	0.0000000	0.988285410	0.01171459	0.0000000	0.00000000	1.0000000

Рис. 2 – Пример степеней принадлежности примеров к значениям атрибутов Fp1 и Fp2 (L – Low, низкий; M – Middle, средний; H – High, высокий).

По итогам данной работы было обнаружено, что выявление различий между классами с  $ассурасу = 1$  возможно.

В Таблице 1 приведена краткая сводка по лучшим результатам классификации. Разбиение по одной зоне указывает на дерево глубиной 1, по одной зоне с последующим разбиением на три – на дерево глубиной 2. Пример нечеткого дерева для показателя Аспр при пробе открытых глаз на дельта-диапазоне представлен на Рисунке 3.

Таблица 1 – Разбиения с ассигуры = 1

Показатель	Аср	Аср	Аmax	Аmax
Диапазон	open	close	open	close
Дельта	Разбиение по Cz Fp1 F3 Fp2	-	-	-
Тета	Разбиение по Fp1 C3 Fp2 C3	Разбиение по C3	Разбиение по T3	-
Альфа	-	Разбиение по Fz	-	-
Бета1	Разбиение по F7	Разбиение по T5	-	Разбиение по F8
Бета2	-	Разбиение по Fp1	-	Разбиение по Fp1 O1 T5 F7
Показатель	Аогиб	Фср	Fmax	Fmax
Диапазон	open	open	open	close
Дельта	Разбиение по Cz	-	Разбиение по T6	-
Тета	-	Разбиение по T5	Разбиение по O2	-
Альфа	-	-	Разбиение по T3 P4 Cz Cz	Разбиение по Pz
Бета1	Разбиение по F7	Разбиение по Fp2 Pz C4 Pz	-	Разбиение по P3 T4 F4 F4
Гамма	-	-	-	Разбиение по O2 F8 T6 T6

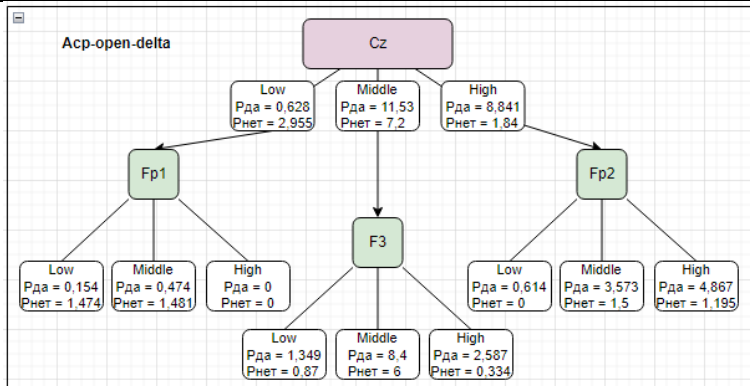


Рис. 3 - Пример нечеткого дерева для показателя Аср при пробе открытых глаз на дельта-диапазоне.

Коэффициент  $R_{да}$  показывает отношение к классу преступников, а  $R_{нет}$  – к классу несудимых. Принадлежность к целевому классу для новой записи находится по формуле 2:

$$\sigma_j = \frac{\sum_l \sum_k P_k^l \cdot \mu_l(D_j) \cdot \chi_k}{\sum_l (\mu_l(D_j) \cdot \sum_k P_k^l)}, \quad (2)$$

где  $P_k^l$  - коэффициент соотношения примеров листа дерева  $l$  для значения целевого класса  $k$ ,  $\mu_l(D_j)$  - степень принадлежности примера к узлу  $l$ ,  $\chi_k$  - принадлежность значения целевого класса  $k$  к положительному значению исхода классификации.

Полученные результаты по примеру из Рисунка 3 для пяти испытуемых при тестировании построенного нечёткого дерева решений представлены в Таблице 2:

**Таблица 2 – Результаты тестирования**

Номер испытуемого	Является преступником на %	Не является преступником на %	Настоящая принадлежность
1	17.53268 %	82.46732 %	Несудимый
2	17.53268 %	82.46732 %	Несудимый
3	61.54374 %	38.45626 %	Преступник
4	78.03823 %	21.96177 %	Преступник
5	60.06248 %	39.93752 %	Преступник

Результаты показали, что метод нечетких деревьев решений может быть применен для решения данной задачи. В будущем планируется увеличить глубину деревьев, добавить новые функции принадлежности, а также исследовать деление данных на другие классы, которые уже были рассмотрены в прошлой работе.

#### Литература:

1. Кутузова И.А. Классификация агрессивного и преступного поведения людей на основе анализа данных ЭЭГ / И.А. Кутузова, Т.В. Авдеенко ; науч. рук. Т. В. Авдеенко. – Наука. Технологии. Инновации : сб. науч. тр. 16 Всерос. науч. конф. молодых ученых, Новосибирск, 5–8 дек. 2022 г. : в 11 ч. – Новосибирск : Изд-во НГТУ, 2022. – Ч. 2. – С. 22–26.
2. Бегенова С.Б. Исследование алгоритма построения нечетких деревьев решений на основе интеграции энтропии и теории нечетких множеств / С.Б. Бегенова, Т. В. Авдеенко. – IV Международная конференция и молодежная школа «Информационные технологии и нанотехнологии» : сб. науч. тр. ИТНТ-2018, Самара, 24-27 апреля 2018 г. – Самара : Изд-во Предприятие "Новая техника", 2018. – С. 2621–2628.



3. Cintra M. E., Monard M. C., Camargo H. A. Fuzzy DT- a fuzzy decision tree algorithm based on C4.5/ M. E. Cintra, M.C. Monard, H.A. Camargo // 2012 CBSF -Brazilian Congress on Fuzzy Systems. – 2012. – pp. 199–211.

4. Janikow C.Z., “Fuzzy Decision Trees: Issues and Methods” / C.Z. Janikow // IEEE Transactions of Man, Systems, Cybernetics. – 1998. – Vol 28, Issue 1.5.

## **НЕЙМАНА ИСЧИСЛЕНИЯ СОЦИОЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ПО ДАННЫМ СТРАТИФИЦИРОВАННЫХ ВЫБОРОК ЕСТЕСТВЕННАЯ МОДИФИКАЦИЯ МЕТОДА**

**И.О. Кутырев**

**Новосибирский государственный университет  
экономики и управления “НИНХ”,  
г. Новосибирск, igkutyrev@gmail.com  
Научный руководитель: Зотьев Д.Б., д.ф.-м.н.**

*В статье рассматривается тема оценки неизвестных социоэкономических показателей по данным стратифицированных случайных выборок. При помощи модификации метода Неймана получены выражения для исчисления оценок, выполнены исследования на основе смоделированных данных.*

*The article describes the topic of estimation of unknown socioeconomic characteristics based on the stratified random samples. Using a modification of the Neumann method, expressions for estimates evaluation are obtained, and studies based on modeled data are performed.*

При проведении социоэкономических исследований зачастую невозможно изучить всю исследуемую совокупность объектов. Это связано как со значительными временными, так и материальными затратами на сбор данных об исследуемых показателях. Поэтому, перед проведением исследования обычно формируют выборку, которая, затем, подвергается исследованию.

Простейшим способом построения выборки является “простая случайная выборка”. Однако, простая случайная выборка требует для каждого объекта равной вероятности быть включенным в выборку, что не всегда возможно. Также, зачастую, исследователи имеют информацию о дополнительных характеристиках исследуемой совокупности, что позволяет построить “стратифицированную случайную выборку”.

Рассмотрим задачу, состоящую в оценке неизвестного значения  $x$  некоторого параметра генеральной совокупности. Стратифицированная случайная выборка реализуется следующим образом: перед формированием выборки генеральная совокупность разбивается на  $k$  “страт”, к каждой из которых применяется метод простой случайной выборки[1].

Пусть, в результате стратификации были получены выборочные показатели параметра  $x$  в стратах:

$$X_{e,1}, X_{e,2}, \dots, X_{e,k}. \quad (1)$$

Тогда, выборочная оценка параметра  $x$ , является линейной комбинацией выборочных оценок этого параметра в стратах

$$X_e = p_1 X_{e,1} + p_2 X_{e,2} + \dots + p_k X_{e,k}, \quad i = 1 \dots k. \quad (2)$$

Вычислим дисперсию такой оценки:

$$\begin{aligned} \sigma^2(X_e) &= E\left((X_e - E(X_e))^2\right) = E(X_e^2) - (E(X_e))^2 = \\ &= \sum_{i=1}^k p_i^2 E(X_{e,i}^2) + 2 \sum_{i < j} p_i p_j E(X_{e,i} X_{e,j}) - \\ &= \sum_{i=1}^k p_i^2 E(X_{e,i})^2 - 2 \sum_{i < j} p_i p_j E(X_{e,i}) E(X_{e,j}). \end{aligned} \quad (3)$$

Так как отбор элементов страт происходит независимо, положим

$$E(X_{e,i} X_{e,j}) = E(X_{e,i}) E(X_{e,j}), \quad i = 1 \dots k, j = 1 \dots k, i \neq j. \quad (4)$$

Тогда

$$\begin{aligned} \sigma^2(X_e) &= \sum_{i=1}^k p_i^2 E(X_{e,i}^2) - \sum_{i=1}^k p_i^2 E(X_{e,i})^2 = \sum_{i=1}^k p_i^2 \sigma^2(X_{e,i}) \\ \sigma(X_e) &= \sqrt{\sum_{i=1}^k p_i^2 \sigma^2(X_{e,i})}. \end{aligned} \quad (5)$$

Согласно методу Неймана, оценка параметра должна обеспечивать минимальное среднеквадратическое отклонение[2]. Следовательно, подбирать коэффициенты линейной комбинации (2) необходимо таким образом, чтобы полученные значения обеспечивали минимизацию

выражения (5). Однако, данный подход не учитывает разницу в объемах страт.

В отличие от метода Неймана, на практике зачастую применяется метод “пропорциональной стратифицированной выборки”, при котором коэффициенты линейной комбинации вычисляются следующим образом:

$$w_i = \frac{N_i}{N}, \quad i = 1 \dots k, \quad (6)$$

где  $N_i$  – количество элементов в страте  $i$ ,  $N$  – количество элементов в генеральной совокупности [3]. Проведем модификацию метода Неймана, в которой учтем разницу в объемах страт. Пусть оценка  $x$  имеет вид

$$X_\sigma = p_1 w_1 X_{\sigma,1} + p_1 w_2 X_{\sigma,2} + \dots + p_1 w_k X_{\sigma,k}. \quad (7)$$

Повторяя преобразования (3) и (5) для (7), получим оценку среднеквадратического отклонения

$$\sigma(X_\sigma) = \sqrt{\sum_{i=1}^k p_i^2 w_i^2 \sigma^2(X_{\sigma,i})}. \quad (8)$$

Воспользуемся методом неопределенных множителей Лагранжа для минимизации  $\sigma(X_\sigma)$  по  $p_i$  [4]:

$$\begin{aligned} \sum_{i=1}^k p_i^2 w_i^2 \sigma^2(X_{\sigma,i}) &\rightarrow \min \\ \sum_{i=1}^k p_i w_i X_{\sigma,i} &= \sum_{i=1}^k w_i X_{\sigma,i}. \end{aligned} \quad (9)$$

Составляя функцию Лагранжа и приравнявая к нулю соответствующие частные производные получим выражения для  $p_i$

$$\left\{ \begin{aligned} p_i &= \frac{\lambda X_{\sigma,i}}{2w_i \sigma(X_{\sigma,i})^2} \\ \lambda &= \frac{\sum_{i=1}^k w_i X_{\sigma,i}}{\sum_{i=1}^k \frac{X_{\sigma,i}^2}{2\sigma(X_{\sigma,i})^2}} \end{aligned} \right., \quad i = 1, \dots, k, \quad (10)$$

которые позволяют найти коэффициенты  $p_i$ , минимизирующие среднеквадратическое отклонение оценки (8).

Для исследования полученных выражений смоделируем несколько независимых страт, формирующих генеральную совокупность. Объем каждой страты – случайное значение из равномерного распределения диапазона [1; 2000000]. Объем выборки из страты – случайное значение из равномерного распределения диапазона [1; объем страты]. Характеристика  $X$  каждого элемента – случайное значение из нормального распределения, параметры которого приведены в таблице 1. На основе выборки вычислим выборочное среднее и выборочное среднеквадратическое отклонение. Таблица 1 показывает смоделированные исходные данные эксперимента.

Таблица 1 – Исходные данные эксперимента

№	объем страты	объем выборки	$X_i$	$X_{e,i}$	$\sigma_i$	$\sigma(X_{e,i})$	$w_i$
1	1641203	316088	0.5	0.5	0.05	0.0495	0.1416
2	1629175	867513	1.0	1.0005	0.1001	0.1002	0.1405
3	1658382	683091	1.5	1.4985	0.1501	0.1483	0.1431
4	1352237	270237	2.0	2.0023	0.1999	0.202	0.1166
5	9479	1712	2.5	2.5063	0.2498	0.252	0.0008
6	1191253	71286	3.0	3.0033	0.3001	0.2991	0.1027
7	616187	176744	3.5	3.4913	0.3498	0.3503	0.0531
8	1302274	443139	4.0	4.0027	0.4002	0.3993	0.1123
9	585742	305194	4.49	4.5013	0.4496	0.4475	0.0505
10	1602503	837569	5.0	5.0081	0.4997	0.4962	0.1382

В результате эксперимента значение  $\sigma(X_e)$  для метода пропорциональной стратифицированной выборки составило **0.0995**, для предлагаемого метода – **0.0799**.

Таким образом, среднеквадратическое отклонение предлагаемого метода оказалось примерно на 20% ниже, чем у метода пропорциональной стратифицированной выборки. Этот результат показывает, что получение выражения (10) можно применять для вычисления оценки неизвестных значений параметров генеральной совокупности по стратифицированной выборке, а классический метод пропорциональной стратифицированной выборки не обеспечивает минимальность среднеквадратического отклонения оценки.

Литература:

1. Новиков С. Стратифицированная выборка в социологическом исследовании // Мониторинг. – 2001. – №4. – URL:

- <https://cyberleninka.ru/article/n/stratifikirovannaya-vyborka-v-sotsiologicheskome-issledovanii> (дата обращения: 01.11.2023).
2. Neyman J. On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection // Journal of the Royal Statistical Society. – 1934. – № 97. – p. 558-625
  3. Чуриков А. Случайные и неслучайные выборки в социологических исследованиях // Социальная реальность. – 2007. – №4. – С. 89-109.
  4. Майорова Н.Л., Глазков Д.В. Методы оптимизации: учебное пособие. – Ярославль: Изд-во ЯрГУ, 2015. – 111 с

## **УСРЕДНЯЮЩАЯ КВАДРАТИЧНАЯ РЕГРЕССИЯ ЭКСПЕРТНЫХ ОЦЕНОК КАЧЕСТВА СМАРТФОНА**

**А.А. Махин**

**Новосибирский государственный университет**

**экономики и управления «НГУЭУ»,**

**г. Новосибирск, [kislik0fist@mail.ru](mailto:kislik0fist@mail.ru)**

**Научный руководитель: Затьев Д.Б., д.ф.-м.н., профессор**

*Статья посвящена проблеме построения рейтингов однородных объектов по качеству с учетом нескольких частных качеств, для чего предлагается использовать агрегирующие, идемпотентные, строго монотонные, сдвиг-инвариантные многочлены (СМ) в соответствии с методикой, изложенной в [4], [5].*

*The article is devoted to the problem of constructing ratings of homogeneous objects by quality, taking into account several particular qualities, for which it is proposed to use aggregating, idempotent, strictly monotonic, shift-invariant polynomials (SM) in accordance with the methodology outlined in [4], [5].*

Многокритериальная оптимизация в практическом аспекте сводится к проблеме выбора наиболее предпочтительного объекта с учетом его различных и полезных свойств. Для ее решения часто используются комплексные показатели качества, при вычислении которых важную роль играют субъективные, экспертные оценки.

При математическом моделировании многокритериальных рейтингов с учетом нескольких частных качеств могут быть полезными строго монотонные, сдвиг-инвариантные, идемпотентные, агрегирующие операторы [2], к которым относятся нормализованные средние функции

(НС), введенные в [3]. Последние определяются, как такие непрерывно-дифференцируемые на  $n$  – мерном кубе  $Q^n = \{q \in \mathbb{R}^n: 0 \leq q_i \leq 1 \forall i = 1, 2, \dots, n\}$  или на всем пространстве  $\mathbb{R}^n$  функции  $q = f(q_1, \dots, q_n)$ , что  $f(0, 0, \dots, 0) = 0$  и для всех  $(q_1, \dots, q_n) \in \mathbb{R}^n$  выполнены условия:

$$\forall i \in \{1, \dots, n\} \quad 0 < q_i < 1 \Rightarrow \sum_{i=1}^n \frac{\partial f}{\partial q_i} = 1 \quad \text{и} \quad \frac{\partial f}{\partial q_i} > 0 \quad (1)$$

$$\forall i \in \{1, \dots, n\}$$

Дифференциальное уравнение в (1) является обобщением условия  $\sum_{i=1}^n w_i = 1$ , которому должны удовлетворять весовые коэффициенты  $w_i > 0$  при наиболее распространенном способе вычисления комплексного показателя [1], как взвешенного среднеарифметического частных показателей:

$$q = \sum_{i=1}^n w_i q_i \quad (2)$$

Среди различных методов подбора весовых коэффициентов  $W_i$  одним из наиболее научно обоснованных является т.н. экспертно-статистический метод [6]. При наличии экспертной оценки  $q = q_{\text{exp}}^j$  комплексного показателя  $q$  для каждого  $j$  – го объекта в некоторой эмпирической совокупности из  $N$  объектов и при известных значениях частных показателей  $q_i = q_i^j$ , где  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, N\}$ , коэффициенты  $w_i$  подбираются так, чтобы аппроксимировать оценки  $q_{\text{exp}}^j$  числами, найденными по формуле (2) значениями. Для этого решается задача оптимизации:

$$\sum_{i=1}^n \left( q_{\text{exp}}^j - \sum_{i=1}^n w_i q_i^j \right)^2 \rightarrow \min \quad \sum_{i=1}^n w_i = 1 \quad \forall i \quad w_i \geq 0 \quad (3)$$

Абсолютная погрешность                      такой аппроксимации                       $\Delta =$

$$\max_{j=1, \dots, N} |q_{\text{exp}}^j - \sum_{i=1}^n w_i q_i^j|.$$

В настоящей работе рассматриваются НСМ  $f(q_1, q_2, q_3, q_4)$  степени 2, выражающий комплексный показатель  $q$ , посредством которого решается практическая задача на моделирование рейтинга смартфонов по четырем качествам – экран, основная камера, фронтальная камера и емкость батареи. В настоящей работе качества оцениваются на основании экспертных рейтингов, публикуемых компанией DXOMARK [6].

Многочлен степени 2 от  $q_1, q_2, q_3, q_4$  выглядит следующим образом:

$$f(q_1, q_2, q_3, q_4) = a_1 q_1^2 + a_2 q_2^2 + a_3 q_3^2 + a_4 q_4^2 + b_{12} q_1 q_2 + b_{13} q_1 q_3 + b_{14} q_1 q_4 + b_{23} q_2 q_3 + b_{24} q_2 q_4 + b_{34} q_3 q_4 + c_1 q_1 + c_2 q_2 + c_3 q_3 + c_4 q_4$$

Варьируя свободные параметры  $b_{ik}$  при  $i < k$  и  $c_l$  при  $l < 4$ , можно аппроксимировать экспертные оценки  $q = q_{exp}^j$ , решая задачу оптимизации:

$$\sum_{j=1}^n (f(q_1^j, q_2^j, q_3^j, q_4^j) - q_{exp}^j)^2 \rightarrow \min$$

С помощью надстройки «Поиск решения» пакета EXCEL было найдено решение  $a_i, b_{ik}, c_l$  задачи (4). В результате были найдены коэффициенты следующего СМ степени 2:

$$f(q) = -0,08q_1^2 - 0,093q_2^2 + 0,16q_1q_2 + 0,025q_2q_4 + 0,16q_1 + 0,327q_2 + 0,513q_4$$

Погрешность, с которой многочлен (5) аппроксимирует экспертные оценки  $q_{exp}^j$  имеет значение  $\Delta = 0,133$ .

Решена также задача (3). В итоге получено  $q = 0,357q_2 + 0,643q_4$ . Погрешность, с которой эта линейная функция аппроксимирует экспертные оценки  $q_{exp}^j$  имеет значение  $\Delta = 0,16$ .

Возрастание погрешности аппроксимации при переходе от квадратичного к линейному СМ составило 20%. Модель экспертного рейтинга, полученная с помощью традиционного подхода (2), является несколько более грубой, чем полученная с помощью СМ (5).

Найденный таким образом СМ  $f(q_1, q_2, \dots, q_n)$  предлагается использовать для вычисления комплексного показателя  $q = f(q_1, \dots, q_n)$  качества произвольного объекта из генеральной совокупности, характеризуемого частными показателями качества  $q_i$ . Тогда общий рейтинг генеральной совокупности определяется отношением предпочтения между объектами, согласно которому большему значению комплексного показателя  $q$  отвечает более предпочтительный объект.

Доказана практическая применимость метода моделирования мультикритериальных рейтингов посредством СМ степеней выше первой. Подробно рассмотрен пример конструирования квадратичного СМ, который значительно увеличивает число подбираемых параметров комплексного показателя качества.

Литература:

1. Азгальдов Г.Г., Райхман Э.П. О квалиметрии. Под ред. Гличева А.В. М.: Издательство стандартов, 1973г. – С. 172.

2. Прикладная статистика: Классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин // М.: Финансы и статистика, 1989. – С. 334, 421-424.

3. *Брызгалин Г.И.* Введение в теорию качества. Волгоград: издательство Волгоградского политехнического института, 1988. - С. 91.

4. *Зотьев Д.Б.* О нормализованных средних критериях, интерполирующих экспертные оценки. Справочник. Инженерный журнал, 2012, 184. - № 7. - С. 50-56.

5. *Зотьев Д.Б.* Нормализованные средние и проблема свертывания показателей качества. Справочник. Инженерный журнал, 146, 2009, № 5, С. 43-48.

6. *Кондрашова Н.В., Данилов И.С., Совершенствование алгоритма расчета интегрального показателя устойчивого развития экономического субъекта.* Современная экономика: проблемы и решения, 2002. - № 5. - С. 54-66.

7. Официальный сайт лаборатории оценки качества «DXOMARK». URL: <https://www.dxomark.com/smartphones/> (дата обращения 02.12.2022).

8. *Goswami S.S., Behera D. K.* Evaluation of the best smartphone model in the market by integrating fuzzy-AHP and PROMETHEE decision-making approach. Decision, 48, 2021. - № 1. - P. 71–96.

## **ВЫЯВЛЕНИЕ РОБАСТНЫХ ЧАСТНЫХ ВЗАИМОСВЯЗЕЙ ПО ДАННЫМ ЭЭГ**

**Д.А. Рыбачёк**

**Новосибирский государственный технический университет,**

**г. Новосибирск, [darierybacyok@gmail.com](mailto:darierybacyok@gmail.com)**

**Научный руководитель: Тимофеева А.Ю., к.э.н.**

*В данной работе рассматривается робастный коэффициент частной корреляции для выявления взаимосвязей между метриками коннективности ЭЭГ в состоянии покоя и показателями вербального и зрительно-пространственного компонентов интеллекта. В результате исследования были отобраны метрики коннективности ЭЭГ, имеющие устойчивую связь с показателями интеллекта.*

*This paper examines the robust partial correlation coefficient to identify correlations between resting-state EEG connectivity metrics and measures of verbal and visual-spatial components of intelligence. As a result of the study, EEG connectivity metrics were selected that have a stable correlation with intelligence components.*



Гипотеза о нейронной эффективности полагает, что люди с высоким уровнем интеллекта при выполнении когнитивных задач демонстрируют более низкую активность мозга по сравнению с людьми, имеющими более низкий уровень интеллекта. Однако исследования, изучающие взаимосвязь активностей мозга в состоянии покоя с показателем интеллекта человека, имеют противоречивые результаты. Решение данной задачи осложняется тем, что метрики коннективности электроэнцефалограмм (ЭЭГ) сильно коррелируют друг с другом, а также обработанные данные ЭЭГ сильно зашумлены. Всё это может привести к выявлению ложных корреляций. Поэтому для выявления устойчивой взаимосвязи было решено попробовать робастный частный коэффициент корреляции.

Целью данной работы является исследование применимости робастного аналога частного коэффициента корреляции для устойчивого выделения взаимосвязей между метриками коннективности ЭЭГ в состоянии покоя и показателями интеллекта человека.

Для оценки синхронизации между парой сигналов в данных ЭЭГ использовалась мнимая когерентность (iМОСН). В качестве порога использовался 70-процентный квантиль показателя мнимой когерентности между ЭЭГ-каналами для каждого человека. Значения iМОСН ниже порога обнулялись. На основе матрицы значений iМОСН построен взвешенный граф [1].

Готовые экспериментальные данные содержат метрики коннективности взвешенного графа, основанные на данных ЭЭГ 107 испытуемых в состоянии покоя: центральность по собственному вектору, модулярность, средняя и характерная длина пути, диаметр, коэффициент кластеризации, центральность по близости. В исследовании используются результаты четырёх субтестов, два из которых вербальные (IQ2, IQ3), и два зрительно-пространственных (IQ7, IQ8).

Коэффициент частной корреляции показывает степень статистической связи двух переменных относительно друг друга после устранения влияния одной или нескольких переменных. В данной работе коэффициент частной корреляции вычисляется на основе коэффициента корреляции Пирсона.

Для расчёта робастного коэффициента частной корреляции воспользуемся критерием LOOT (leave-one-out test, критерий поэлементного удаления) [1]. Процедура поэлементного удаления может повторяться  $n$  раз для набора данных, состоящего из  $n$  точек: из данных выбрасывается одна точка и считается коэффициент корреляции без удалённой точки. Таким образом, в результате работы критерия имеется  $n$  коэффициентов корреляции. На основе средневзвешенного значения

полученных коэффициентов рассчитывается робастный коэффициент корреляции. Взвешиваем коэффициенты корреляции в соотношении их отличия от актуального коэффициента корреляции:

$$w_i = \left| r_{actual} - r_{i, LOOT} \right|^a \quad (1)$$

где  $r_{actual}$  – значение коэффициента корреляции, рассчитанное по всей выборке,  $r_{i, LOOT}$  – значение коэффициента корреляции при удалении  $i$ -й точки данных в LOOT,  $a = 1 + n/12$ .

Тогда робастный коэффициент частной корреляции рассчитывается по формуле:

$$r_{Rob} = \frac{\sum_i^n w_i r_{i, LOOT}}{\sum_i^n w_i}. \quad (2)$$

При помощи перестановочного критерия можно проверить значимость робастного аналога коэффициента корреляции. Так как выборка содержит небольшое количество наблюдений (107), было решено взять уровень значимости  $\alpha = 0.1$ . Для удобства проводились случайные перестановки откликов (IQ2, IQ3, IQ7, IQ8) [2].

Для сравнения также были подсчитаны коэффициенты корреляции Пирсона и Спирмена. В таблице 1 приведены полученные коэффициенты корреляции между метриками ЭЭГ в тэта диапазоне и показателем интеллекта IQ8 (зрительно-пространственный). Коэффициенты, значимые на уровне 5% помечены символами \*\*, на уровне 10% – символом \*.

**Таблица 1 – Значения коэффициентов корреляции**

Метрика	Пирсон	Спирмен	Робастный частный
Центральность по собственному вектору	0.06849	0.11578	-0.01305
Модулярность	0.12045	0.12116	0.09945
Средняя длина пути	0.17759*	0.22708*	0.03244
Характерная длина пути	0.17574*	0.22568*	0.12291
Диаметр	0.20947**	0.23688*	0.21223*
Коэффициент кластеризации	-0.06992	-0.04874	-0.10538
Центральность по близости	-0.13872	-0.17356*	0.17238*

При помощи коэффициента корреляции Пирсона и Спирмена удалось выявить следующие значимые корреляции: положительная связь IQ8 со средней длиной пути, характерной длиной пути и диаметром. Также при помощи коэффициента корреляции Спирмена удалось выделить отрицательную взаимосвязь с центральностью по близости. Робастный аналог частного коэффициента корреляции выделил положительную взаимосвязь с диаметром и отрицательную – с центральностью по близости.

По данным в диапазоне тэта и показателем интеллекта IQ8, все коэффициенты корреляции выявили положительную связь с диаметром.

Сопоставив результаты по всем данным, можно заметить, что по некоторым метрикам были получены значимые (на уровне 10%) коэффициенты корреляции как при помощи коэффициента корреляции Спирмена, так и при помощи робастного коэффициента частной корреляции. Метрики и значения коэффициентов приведены в таблице 2.

**Таблица 2 – Значимые корреляции между показателями интеллекта и метриками коннективности**

Метрика	Отклик	Спирмен	Роб. частный
Модулярность	IQ2 дельта	0.19347	0.15002
Диаметр	IQ7 альфа	-0.1752	-0.1307
Модулярность	IQ7 тэта	-0.17365	-0.1792
Диаметр	IQ8 тэта	0.23688	0.21223

Таким образом, можно утверждать, что данные метрики действительно имеют устойчивую взаимосвязь с показателями интеллекта.

В данной работе показано, что робастный коэффициент частной корреляции может использоваться для отбора наиболее информативных признаков, которые влияют на отклик сами, а не за счёт взаимосвязей между собой.

#### Литература:

1. Выявление устойчивых связей между показателями коннективности ЭЭГ и компонентами интеллекта / Т.В. Авдеенко, А.Ю. Тимофеева, М.Ш. Муртазина, О.М. Разумникова // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2022. – Т. 11, №4. – С. 19-36.
2. Legendre P. Comparison of permutation methods for the partial correlation and partial mantel tests. // Journal of Statistical Computation and Simulation. – 2000. – Vol. 67, iss. 1. – P. 37–73.

## ФОРМУЛЫ ДЛЯ ЭФФЕКТИВНОГО ПОРТФЕЛЯ АКЦИЙ В СТАНДАРТНОЙ МОДЕЛИ МАРКОВИЦА

**В.Е. Соколов**

**Новосибирский государственный университет  
экономики и управления «НИНХ»,**

**г. Новосибирск, slava.sokolov1999@mail.ru**

**Научный руководитель: Зотьев Д.Б., д. ф.-м. н., профессор каф.  
математики и ЕН НГУЭУ**

*Доклад посвящен эмпирической проверке явных формул для параметров эффективного портфеля акций, полученных на основании теории Марковица. Использованы данные о котировках акций «Газпром», «Московская Биржа», «Сбербанк России» и «Светофор» во временном интервале с января по сентябрь 2023 года. Выяснилось, что относительная погрешность волатильности эффективного портфеля из указанных выше акций, полученного посредством оптимизации стандартной модели Марковица, не превышает 0,27%. Таким образом, подтверждена справедливость явных формул.*

*The report is devoted to empirical testing of explicit formulas for the parameters of an effective stock portfolio obtained based on the Markowitz theory. Data were used on stock quotes of Gazprom, the Moscow Exchange, Sberbank of Russia and Svetofor in the time interval from January to September 2023. It turned out that the relative volatility error of the effective portfolio of the above stocks, obtained by optimizing the standard Markowitz model, does not exceed 0.27%. Thus, the validity of the explicit formulas is confirmed.*

При инвестировании в акции или иные волатильные активы инвестор рассчитывает на получение максимальной прибыли и стремится минимизировать риск убытков. Достижение этих двух целей одновременно невозможно, т.к. они противоречат друг другу. Средством для достижения компромисса является диверсифицированный портфель акций. Данный термин обозначает некоторое распределение инвестиционного капитала между акциями различных эмитентов, так что на покупку акций каждого эмитента из представленных в «портфеле» расходуется определенная доля от общего инвестиционного капитала. Эти доли являются параметрами портфеля, которые следует оптимизировать. Точное определение такой оптимизации было дано в 1952, когда сотрудник Rand Corporation Гарри Марковиц опубликовал статью [1]. В ней он определил риск инвестора, как среднее квадратическое отклонение доходности и предложил считать оптимальным тот портфель, который при желаемой доходности имеет минимальный риск.

Для некоторого числа  $k > 1$  обозначим  $X_1, X_2, \dots, X_k$  случайные величины, принимающие значения доходностей вложения в акции эмитентов № 1, 2, ...,  $k$ . Обозначим  $p_1, p_2, \dots, p_k$  доли инвестируемой суммы, вкладываемые в акции № 1, 2, ...,  $k$  соответственно. Тогда доходность портфеля выражается суммой произведений доходностей каждой акции на её долю в портфеле.

Риск портфеля рассчитывается из формулы:

$$X = \sum_{i=1}^k p_i^2 \sigma_i^2 + 2 \sum_{i < j} p_i p_j \sigma_i \sigma_j r_{ij} \quad (1)$$

г  
д

е  $r_{ij}$  – коэффициент корреляции между  $X_i$  и  $X_j$ . Предположим, что ожидаемая доходность портфеля равна некоторому  $P$ , т.е.

$$p_1 + p_2 + \dots + p_k = 1 \quad (3)$$

Таким образом для того, чтобы при заданной доходности портфеля  $P$  минимизировать риск нужно решить задачу оптимизации:

$$X = \sum_{i=1}^k p_i^2 \sigma_i^2 + 2 \sum_{i < j} p_i p_j \sigma_i \sigma_j r_{ij} \rightarrow \min \quad (4)$$

п  
р По Марковицу эффективным портфелям отвечают такие точки  $\sigma, P$ , что нельзя ни уменьшить риск  $\sigma$  без уменьшения доходности  $P$  ни

у  
в Согласно Д.Б. Зотьеву, параметры эффективного портфеля выражаются следующими формулами:

$$p_i = A_i P + B_i \quad (5)$$

$$W_i = \beta_i \sigma_i^2 = 1/k \alpha_j \sigma_j - \alpha_i \sigma_i = 1/k \beta_j \sigma_j = 1/k \alpha_j \sigma_j - j = 1/k \beta_j \sigma_j = 1/k \alpha_j \sigma_j \quad (6)$$

$$B_i = \alpha_i \sigma_i^2 = 1/k \beta_j \sigma_j - \beta_i \sigma_i = 1/k \alpha_j \sigma_j - j = 1/k \beta_j \sigma_j - j = 1/k \alpha_j \sigma_j \quad (7)$$

$$A_j \sigma_j^2 = 1/k \alpha_j \sigma_j - \alpha_j \sigma_j = 1/k \beta_j \sigma_j - j = 1/k \alpha_j \sigma_j \quad (6)$$

$$B_j \sigma_j^2 = 1/k \alpha_j \sigma_j - \alpha_j \sigma_j = 1/k \beta_j \sigma_j - j = 1/k \alpha_j \sigma_j \quad (6)$$

$$A_1 + r_{12} \alpha_2 + r_{13} \alpha_3 + \dots + r_{1k} \alpha_k = x_1 \sigma_1 r_{21} \alpha_1 + \alpha_2 + r_{23} \alpha_3 + \dots + r_{2k} \alpha_k = x_2 \sigma_2 \dots \dots \dots r_k$$

$$B_1 \sigma_1^2 = 1/k \alpha_1 \sigma_1 - \alpha_1 \sigma_1 = 1/k \beta_1 \sigma_1 - 1 = 1/k \alpha_1 \sigma_1 \quad (6)$$

$$A_2 + r_{21} \alpha_1 + r_{23} \alpha_3 + \dots + r_{2k} \alpha_k = 1 \sigma_2 r_{12} \beta_1 + \beta_2 + r_{23} \beta_3 + \dots + r_{2k} \beta_k - \beta_2 - 1 + r_{2k} \beta_k = 1 \sigma_2 \dots \dots \dots r_k \beta_1$$

$\beta_j \sigma_j^2$  являются решениями систем уравнений (8) и (9):

$$B_j \sigma_j^2 = 1/k \alpha_j \sigma_j - \alpha_j \sigma_j = 1/k \beta_j \sigma_j - j = 1/k \alpha_j \sigma_j \quad (6) \quad 124$$

$$A_j \sigma_j^2 = 1/k \alpha_j \sigma_j - \alpha_j \sigma_j = 1/k \beta_j \sigma_j - j = 1/k \alpha_j \sigma_j \quad (6)$$

$$rk2\beta2+rk3\beta3+\dots+rk,k-1\beta k-1+\beta k=1\sigma k \quad (9)$$

Минимальный риск при этом выражается формулой:

$$AP^2+2BP+C \quad (10)$$

где коэффициенты А, В и С рассчитываются из формул (11),(12) и (13):

$$i=1k\sigma i2Ai2+2i<jkrij\sigma i\sigma jAiAj \quad (11)$$

$$i=1k\sigma i2AiBi+i<jkrij\sigma i\sigma j(AiBj+AjBi) \quad (12)$$

$$i=1k\sigma i2Bi2+2i<jkrij\sigma i\sigma jBiBj \quad (13)$$

Для проведения эмпирического прогноза были отобраны акции четырёх Российских компаний, торгуемых на Московской фондовой бирже. Данные о котировках акций были взяты с официального сайта инвестиционного холдинга «Финам». В работе рассмотрены акции таких компаний как ПАО «Газпром», ПАО «Московская Биржа», ПАО «Сбербанк России» и торговая сеть «Светофор». Временной интервал представлен девятью месяцами с января по сентябрь 2023 года.

В результате эмпирического построения и построения на основе явных формул получаем две дуги гиперболы с кривыми Марковица.

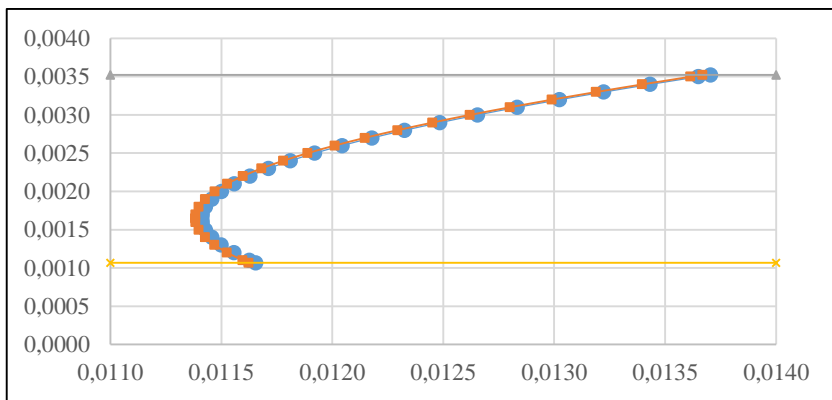


Рис. 2 – Кривые Марковица

Степень отклонения теоретических значений от эмпирических оценивается с помощью формулы:

$$\Delta\sigma = \frac{|\sigma_P - \sigma_T|}{\sigma_T} * 100\% \quad (14)$$

Отклонение на всей построенной кривой не превышает 0,27%, что является незначительной погрешностью.

Таким образом в исследовании были эмпирически проверены и подтверждены теоретические формулы, которые в явном виде дают зависимость долей активов в инвестиционном портфеле от установленной желаемой доходности. Исследование имеет практическую значимость, позволяя инвесторам быстрее и надёжнее просчитывать свой уровень диверсификации портфеля и, как следствие, размещать свои средства с меньшим риском убытков не снижая уровень финансовой отдачи.

Литература:

1. Markowitz H. Portfolio Selection // The Journal of Finance. – 1952. – Vol. 7, N 1. – P. 77-91.
2. Markowitz H., Todd P. Mean-Variance Analysis in Portfolio Choice and Capital Markets. 1987. – 408 p.
3. Финам.ру – финансовый портал: [Электронный ресурс]. URL: <https://www.finam.ru/profile/moex-akcii/gazprom/export/> (Дата обращения: 08.11.2023).

## **СТАТИСТИЧЕСКИЙ АНАЛИЗ УРОВНЯ РАЗВИТИЯ ЗАРЯДНОЙ ИНФРАСТРУКТУРЫ**

**М.Р. Табакаева**

**ФГБОУ ВО «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева»,**

**г. Красноярск, [tabakaevamaria@mail.ru](mailto:tabakaevamaria@mail.ru)**

**Научный руководитель: Федорова Н.В., к.э.н., доцент**

*Одним из ключевых факторов роста электромобильной отрасли является зарядная инфраструктура. Для стимулирования перехода на экологически чистый вид транспорта необходимо продвижение и увеличение количества установленных зарядных станций, их доступности и скорости зарядки. Развитая зарядная инфраструктура способствует устойчивому развитию и прогрессу в области чистой энергетики.*

*One of the key drivers of growth in the electric vehicle industry is charging infrastructure. Promoting and increasing the number of installed charging stations, their availability and charging speed is necessary to stimulate a shift to clean transportation. Developed charging infrastructure promotes sustainable development and progress in clean energy.*

На сегодняшний день развитые и развивающиеся страны признают значимость электромобильной отрасли для снижения выбросов парниковых газов, улучшения качества воздуха и диверсификации энергического сектора, в связи с чем целенаправленная политика многих стран по трансформации глобальной экономики направлена на альтернативный топливно-энергетический баланс. За последние годы произошли фундаментальные изменения на транспортном рынке в Российской Федерации. Правительство Российской Федерации осуществляет политику стимулирования развития зарядной инфраструктуры и электрокаров. В августе 2021 г. Правительство Российской Федерации утвердило концепцию развития производства и использования электротранспорта в Российской Федерации до 2030 г. Концепция по развитию производства и использования электрического автомобильного транспорта на период до 2030 года разработана в соответствии с Федеральным законом «О стратегическом планировании в Российской Федерации»[1].

Распоряжением Правительства Российской Федерации №469-р от 10.03.2022 г. утвержден межбюджетный трансфер из федерального бюджета бюджетам субъектов Российской Федерации в целях реализации мероприятий по развитию зарядной инфраструктуры, а также перечень территорий, где в рамках пилотного эксперимента для создания зарядной инфраструктуры электротранспортных средств на сегодняшний день включено 8 субъектов Российской Федерации и дорог федерального значения. Утвержденная дорожная карта планируется актуализироваться с привлечением широкого круга представителей экспертного сообщества, производителей и потребителей.

В соответствии с распоряжением № 469-р от 10.03.2022 г. целевыми показателями реализации развития зарядной инфраструктуры на 2023 – 2024 гг. является запуск в эксплуатацию 528 быстрых зарядных станций, затраты на закупку оборудования которых возмещаются получателям средств [2].

Всего в России по данным 2chargers.net 4367 точек заряда, из которых 3679 – точки медленной зарядки, 1259 – быстрой. В 2022 г. в России было установлено 688 быстрых электроразрядных станций (ЭЗС) [3].

В настоящее время лидерами по станциям переменного и постоянного тока является Москва (Московская область). Наиболее оснащенные зарядной инфраструктурой регионы страны представлены на рисунке 1 и рисунке 2 [4].

На сегодняшний день центральным проектом развития электромобильной инфраструктуры в Российской Федерации выступают ПАО «Россети». Компания реализует программу по развитию зарядной



инфраструктуры на территории всей страны совместно с Минэнерго России, имея своей целью постепенное распространение зарядных станций в каждом из регионов страны, что в свою очередь стимулирует рост электромобилей.



Рис. 1 – Рейтинг регионов по количеству зарядных станций переменного тока



Рис. 2 – Рейтинг регионов по количеству зарядных станций постоянного тока

Таким образом, в Российской Федерации стремительно развивается зарядная инфраструктура. Реализация различных программ и мер поддержки позволит обеспечить наличие зарядных станций по всей

территории страны, что в свою очередь создаст комфортное использование электромобилей на больших расстояниях и в отдельных регионах страны.

Литература:

1. Распоряжение Правительства РФ от 23.08.2021 №2290-р «Об утверждении Концепции по развитию производства и использования электрического автомобильного транспорта в Российской Федерации на период до 2030 года» // Собрание законодательства РФ. – 30.08.2021. - №35. – Ст. 6327.

2. Постановление Правительства РФ от 10.03.2022 № 469-р «О распределении в 2022 году иных межбюджетных трансфертов из федерального бюджета бюджетам субъектов Российской Федерации в целях реализации мероприятий по развитию зарядной инфраструктуры для электромобилей» // Собрание законодательства РФ. – 21.03.2022. - №12. – Ст. 1889.

3. В России насчитывается более 4 тысяч зарядных станций [Электронный ресурс]. – Режим доступа: URL: <https://www.autostat.ru/news/53199/> (дата обращения: 17.11.2023).

4. В рамках форума «IT-Революция в автобизнесе – 2022» прошла выставка электромобилей [Электронный ресурс]. – Режим доступа: URL: [https://dzen.ru/a/Y3\\_kQtpXOxDo56wr](https://dzen.ru/a/Y3_kQtpXOxDo56wr) (дата обращения: 13.11.2023).

### **КЛАСТЕРИЗАЦИЯ ПАРТИЙ ПОЛУФАБРИКАТОВ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

**Л.Е. Теников, Л.Н. Бодрякова**

**г. Норильск, [isit.zgu.2022@gmail.com](mailto:isit.zgu.2022@gmail.com)**

**Научный руководитель: Бодрякова Л.Н., к.т.н., доцент**

*В статье рассмотрены варианты кластеризации партий полуфабрикатов с применением нейронных сетей. Кластеризация по размерным признакам и цвету с применением самоорганизующихся карт Кохонена. Рассмотрены алгоритмы решения задачи распознавания изображений: контурный анализ, поиск шаблона и сопоставление по ключевым точкам*

*The article discusses options for clustering batches of semi-finished products using neural networks. Clustering by dimensional features and color using self-organizing Kohonen maps. Algorithms for solving the problem of image recognition are considered: contour analysis, pattern search and matching by key points*

Сортировка партий полуфабрикатов является одной из распространенных задач во многих технологических процессах и

различается по количеству учитываемых факторов и сложности задачи. И если партии полуфабрикатов, полученных в результате производственной деятельности, характеризуются высокой однородностью свойств и основная задача – контроль качества. То полуфабрикаты, являющиеся природными материалами, отличаются высокой разнообразностью свойств. К данному направлению относится сортировка шкурок пушно-мехового полуфабриката в процессах скорняжного производства. При ее выполнении учитывается большое количество факторов, свойства отличаются высокой неоднородностью, отсутствуют инструментальные методы их определения и четкий алгоритм выполнения сортировки. Вышеперечисленные факторы определяют трудноформализуемость задачи.

Анализ существующих технических решений при решении аналогичных задач позволил установить, что наиболее перспективным решением является применение технологий искусственного интеллекта, в частности, технологий компьютерного зрения. Данное решение имеет следующие преимущества: возможность комплексной оценки характеристик свойств пушно-мехового полуфабриката, автоматическая кластеризация шкурок, что позволяет автоматизировать сортировку партии на однородные группы.

Авторами выполнены исследования применения алгоритма карт Кохонена для анализа и кластеризации шкурок норки по размерам. Геометрические критерии приведены на рисунке 1. В исследовании использовались данные шести партий шкурок, каждая из которых включала в себя 50 единиц. Алгоритм на входе принимает вектор свойств для каждой шкурки, который включает в себя показатели длины и ширины как в верхней, так и в нижней части шкурки (рисунок 1).

Использование самоорганизующихся карт Кохонена позволило выполнить кластеризацию данных на основе размерных характеристик шкурок. В результате алгоритм выделил пять подмножеств, демонстрирующих сходство в линейных размерах шкурок внутри каждой группы. Значительным является тот факт, что результаты кластеризации, выполненной с помощью алгоритма Кохонена, находятся в соответствии с сортировкой, проведенной экспертами. Характеристика кластеров представлена на рисунке 1.



Рис. 1 – Характеристика кластеров  
 а – учитываемые геометрические характеристики, б – результат кластеризации шкурок

Для сортировки однородноокрашенных шкурок разработано программное обеспечение, позволяющее на основе анализа цифрового изображения шкурки идентифицировать цвет. Задача решается с помощью библиотеки компьютерного зрения OpenCV, после чего данные кластеризуют с помощью нейронных сетей. Машинное обучение выполнено посредством метода K-средних. Тестирование выполнено на партии шкурок каракулево-смушковой группы объемом 50 шт. Апробация дала положительный результат и подтвердила возможность применения технологий в технологических процессах.

Направление дальнейших исследований – разработка методики оценки комплексного показателя цвета шкурок естественной окраски и рисунок топографических участков (рисунок 2). Для решения поставленной задачи одним из наиболее ответственных этапов является распознавание контура изображения.

Рассмотрены наиболее распространенные методы решения задач распознавания изображений: контурный анализ, поиск шаблона и сопоставление по ключевым точкам [1-3].



Рис. 2:

а – неоднородноокрашенная шкурка, б – результат обработки изображения с применением нейронных сетей

Алгоритм **контурного анализа** представляет из себя метод описания, хранения, распознавания, сравнения и поиска графических образов по их контурам. Контур – кривая, описывающая границу искомого объекта. Использование данного подхода предполагает, что контур содержит достаточно информации о форме объекта, при этом внутренние точки не учитываются. Переход к пространству контуров существенно снижает сложность алгоритмов и вычислений, необходимых для обработки изображений. Существенной особенностью контурного анализа является инвариантность относительно смещения контура на тестируемом изображении, вращения и масштаба. Метод отлично подойдет для поиска объекта некоторой заданной формы [1, 2].

При всех достоинствах алгоритма, следует иметь в виду, что область его применения сильно ограничена. Прежде всего, они вызваны проблемами выделения контура на изображении. Например, объект может не иметь чёткой границы, при одинаковой яркости с фоном; невозможно выделить контур из-за зашумленности данных. Контур может быть определен ошибочно в том случае, если искомый объект перекрыт посторонним объектом.

Таким образом, контурный анализ обладает слабой устойчивостью к помехам и при любом нарушении целостности контура или плохой видимости объекта приводят либо к невозможности детектирования, либо к ложным срабатываниям.

Метод **поиска шаблона** используют, если необходимо искать участки изображений, которые максимально схожи с некоторым определенным шаблоном [1, 3]. Входными параметрами метода являются:

- изображение, на котором производится поиск шаблона;
- изображение искомого объекта; при этом размер шаблона обязательно должен быть меньше размера изображения, на котором проводится поиск.

Алгоритм должен найти на тестируемой картинке область, которая лучше всего совпадает с шаблоном. Поиск шаблона производится путем последовательного перемещения его на один пиксель за раз по тестируемому изображению, и оценкой схожести каждой новой области с шаблоном. По результатам проверки выбирается та область, с максимальным коэффициентом совпадения. По сути – это процент совпадения области картинки и шаблона. Данный метод дает хорошие результаты, когда необходимо быстро проверить наличие некоторого объекта на изображении (поиске изображений лиц людей). Его можно адаптировать для поиска шаблона сразу на видеопотоке веб-камеры или камеры мобильного устройства. Недостатки метода: невозможно с уверенностью сказать обнаружен исходный объект или нет, в виду того, что это вероятностная характеристика и зависит от масштаба, углов обзора, поворотов картинки и наличия помех. Вдобавок возможны ложноположительные срабатывания алгоритма, когда имеются общие детали у шаблона и области на тестируемом изображении.

#### Литература:

1. Род Стивенс. Алгоритмы. Теория и практическое применение. – Москва : Эксмо, 2021. – 544 с.
2. What Canny Edge Detection algorithm is all about? // Satyajit Manitra. Medium. URL: <https://medium.com/@ssatyajitmaitra/what-canny-edge-detection-algorithm-is-all-about-103d94553d21> (дата обращения: 27.06.2022).
3. HOG + SVM обнаружение пешеходов // Русские Блоги. URL: <https://russianblogs.com/article/1298987091/> (дата обращения: 27.06.2022).
4. Постолит А. В. Основы искусственного интеллекта в примерах на Python. Санкт-Петербург, 2021. – 448 с.

## ОПТИМИЗАЦИЯ ГЛАДКОСТИ СОПРЯЖЕНИЯ ДВУХ КРИВЫХ ДУГОЙ ОКРУЖНОСТИ

Д.А. Чагочкина

Новосибирский государственный технический университет,  
г. Новосибирск, chagochkina.2020@stud.nstu.ru  
Научный руководитель: Карманов В.С., к.т.н., доцент

*Предложен метод получения сопряженной кривой в виде дуги окружности для таблично заданных кривых. Решение, основанное на системе нелинейных уравнений, учитывает координаты точек на кривых и радиус требуемой окружности. Применение предложенного подхода на модельных данных демонстрирует возможность выбора оптимальных точек для более эффективного построения сопряженной кривой.*

*A method for obtaining a conjugate curve in the form of a circular arc for tabulated curves is proposed. The solution based on a system of nonlinear equations considers the coordinates of points on curves and the radius of the required circle. The application of the proposed approach on model data demonstrates the possibility of choosing optimal points for more efficient construction of the conjugate curve.*

### Введение

Сопряжение кривых находит широкое применение в различных сферах деятельности человека: в машинной графике, робототехнике, проектировании автомобильных и железнодорожных путей и др. Например, в компьютерной графике для создания плавных изгибов и анимаций часто используются сплайны, в частности кривые Безье [1,2], описываемые параметрическими уравнениями, которые используют контрольные точки для определения формы кривой. В проектировании транспортной инфраструктуры большую роль играет применение переходных кривых [3,4], которые позволяют сглаживать переходы между участками с различной кривизной пути и обеспечивает более естественное и плавное движение.

В ходе проектирования в различных областях деятельности человека возникает необходимость сопряжения кривых с окружностью заданного радиуса, что обусловлено техническими требованиями. Для описания сопряженной кривой используются табличные данные дуг кривизны, представленные в виде  $L_1 = (X_1, \dots, X_m)^T$ ,  $L_2 = (Y_1, \dots, Y_n)^T$ , где  $X_i, Y_j$  – точки на кривых  $L_1$  и  $L_2$  соответственно,  $i = \overline{1, m}$ ,  $j = \overline{1, n}$ .

### Постановка задачи

Даны координаты двух точек на двух кривых  $P_1 = (x_1, y_1)$  и  $P_2 = (x_2, y_2)$ , и радиус дуги требуемой окружности  $R$ . Требуется найти координаты центра требуемой окружности  $P_0 = (x_0, y_0)$ , которая будет проходить через заданные две точки с заданным радиусом.

### Метод решения

Решение найдем, подставляя в уравнение окружности значения координат точек  $P_1$  и  $P_2$ :

$$(x - x_0)^2 + (y - y_0)^2 = R^2. \quad (1)$$

Получим нелинейную систему из двух уравнений относительно двух неизвестных  $x_0, y_0$ :

$$\begin{cases} (x_1 - x_0)^2 + (y_1 - y_0)^2 = R^2 \\ (x_2 - x_0)^2 + (y_2 - y_0)^2 = R^2. \end{cases} \quad (2)$$

Решая систему, получим искомые  $x_0$  и  $y_0$ :

$$x_0 = \frac{x_1 + x_2}{2} \pm \frac{y_2 - y_1}{2} \sqrt{4 \frac{R^2}{d^2} - 1}, \quad (3)$$

$$y_0 = \frac{y_1 + y_2}{2} \pm \frac{x_2 - x_1}{2} \sqrt{4 \frac{R^2}{d^2} - 1}. \quad (4)$$

где  $d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$  – квадрат расстояния между заданными точками.

Система не имеет решений, если  $d > 2R$ . Если  $d \leq 2R$ , то система имеет ровно два решения, в зависимости от направления касательных прямых в точках сопряжения.

После получения центра окружности, координаты точек окружности можно найти по формуле (параметрическое задание окружности):

$$\begin{cases} x(t) = x_0 + R \cos t. \\ y(t) = y_0 + R \sin t. \end{cases} \quad (5)$$

где  $t \in [0, 2\pi)$  – параметр.

При задании точек дуги необходимо найти границы изменения параметра  $t \in [t_1, t_2]$  по формуле:

$$t_1 = \operatorname{atan2} \left( \frac{y_1 - y_0}{R}, \frac{x_1 - x_0}{R} \right), \quad (6)$$

$$t_2 = \operatorname{atan2} \left( \frac{y_2 - y_0}{R}, \frac{x_2 - x_0}{R} \right), \quad (7)$$

где  $\operatorname{atan2}(y, x)$  – функция «арктангенс двух аргументов», определяющая угол  $\theta \in (-\pi, \pi]$ , отсчитываемый против часовой стрелки от положительного направления оси  $X$  до точки  $(x, y)$ , с учетом координатной четверти, в которой лежит эта точка.

### Результаты исследований

Рассмотрим применение предложенного подхода на модельных данных. Пусть даны две кривые  $L_1$  и  $L_2$  с заданным набором точек  $(X_1, \dots, X_m)^T$  и  $(Y_1, \dots, Y_n)^T$  соответственно,  $i = \overline{1, m}$ ,  $j = \overline{1, n}$ ,  $m = 258$ ,



$n = 104$ . Для построения кривой путем сопряжения кривых  $L_1$  и  $L_2$  выберем точки  $P_1$  и  $P_2$ , затем проведем к обеим кривым касательные, точка пересечения которых позволяет определить, где лежит центр окружности  $C_0$  (рис. 1).

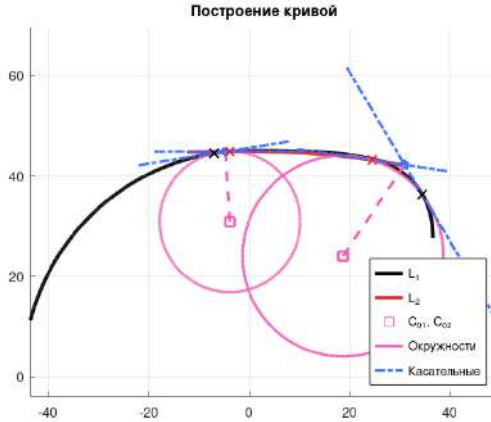


Рис.1 – Пример сопряжения двух кривых дугой окружности

Возникают случаи, когда путем сопряжения двух кривых не получается оптимальной кривой, так как степень кривизны высокая, если фиксировать две точки  $P_1$  и  $P_2$ . Для решения проблемы предлагается следующий подход: на кривой  $L_1$  фиксируют  $P_1$  и перебором точек ищется точка  $P_2$  на кривой  $L_2$ , при этом  $P_2$  ищется в окрестности точки пересечения кривых (рис. 2).

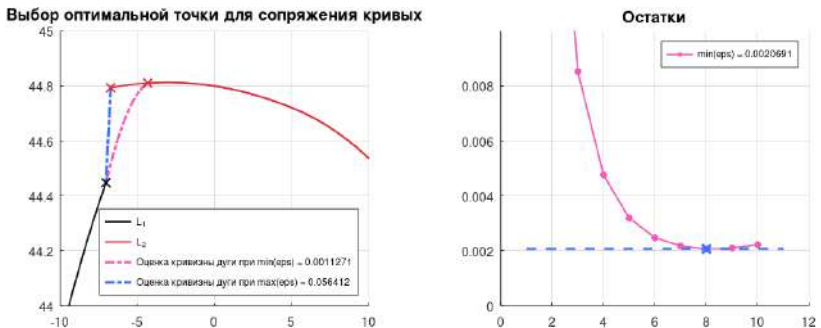


Рис.2 – Выбор оптимальной точки для сопряжения кривых и график остатков

Для нахождения оптимальной кривой оценивается сумма квадратов отклонений ординат точек:

$$\varepsilon = (y_3 - \widehat{y}_3)^2 + (y_4 - \widehat{y}_4)^2 \rightarrow \min,$$

где  $y_3$  – ордината точки  $(x_3, y_3)$ , предшествующей  $P_1$ ,  $\widehat{y}_3$  – ордината точки  $(x_3, \widehat{y}_3)$  на окружности;  $y_4$  – ордината точки  $(x_4, y_4)$ , следующей за оцениваемой точкой  $P_2$ ,  $\widehat{y}_4$  – ордината точки  $(x_4, \widehat{y}_4)$  на окружности. На рис. 3 показана построенная кривая.

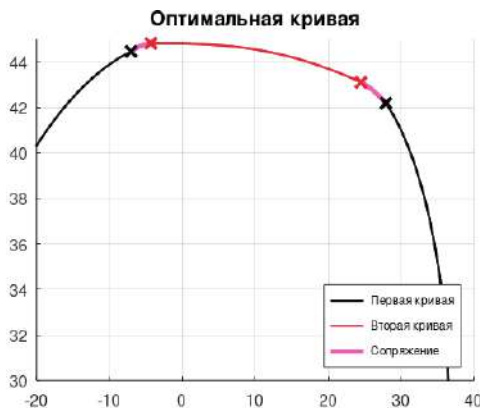


Рис.3 – Построение оптимальной кривой путем сопряжения двух кривых

Таким образом, с помощью предлагаемого подхода можно построить кривую, состоящую из данных кривых и дуги окружности, которая будет оптимизирована по гладкости.

### Выводы и заключение

Рассмотренный метод сопряжения кривых с использованием окружности заданного радиуса представляет эффективный способ создания плавных переходов между точками на кривых, заданных таблично. Результаты исследований позволяют оптимизировать выбор точек сопряжения, обеспечивая создание более естественных и гармоничных кривых при решении задач в различных областях.

### Литература:

1. Елугачев, П. А. Пространственное трассирование автомобильных дорог кривыми Безье: специальность 05.23.11 "Проектирование и строительство дорог, метрополитенов, аэродромов, мостов и транспортных тоннелей" : автореферат диссертации на соискание ученой степени кандидата технических наук / Елугачев Павел Александрович. – Москва, 2007. – 17 с. – EDN NJDHQV.

2. Роджерс Д., Адамс Дж. Математические основы машинной графики. — М.: Мир, 2001.

3. Аникеев Е. А. Алгоритм построения сопряжения криволинейных участков автодороги // Программные продукты и системы. 2008. №2.

4. Autodesk Civil 3D: [Электронный ресурс] // Определения переходной кривой. URL: <http://docs.autodesk.com/CIV3D/2012/RUS/filesCUG/GUID-581518E0-DE06-482E-840C-B997C359048-1286.htm>. (Дата обращения: 22.11.2023)

## **СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭКОНОМИЧЕСКИХ СИСТЕМ СИБИРИ И СКАНДИНАВИИ С ИСПОЛЬЗОВАНИЕМ ГРАФОВ**

**А.А. Шубин**

**Новосибирский государственный технический университет,**

**г. Новосибирск, Andrej-shubin-1997@mail.ru**

**Научный руководитель: Тимофеева А.Ю., к.э.н**

*В данной работе производится сопоставление Сибирского и Скандинавского регионов путём построения графов на основе социально-экономических данных на примере данных государственной статистики. Исследованы две метрики расстояния (Жаккара и на основе коэффициента корреляции Пирсона), построены матрицы смежности, вычислены характеристики взвешенных графов. Для упрощения графа задавалось пороговое значение расстояния. В результате работы выявлено, что полученные характеристики говорят о большей центральности и связности графа Нордического региона в сравнении с графом Сибирского региона.*

*This paper compares Siberian and Scandinavian regions by constructing graphs on the basis of socio-economic data on the example of state statistics data. Two distance metrics (Jaccard and on the basis of Pearson correlation coefficient) are investigated, adjacency matrices are constructed, and characteristics of weighted graphs are calculated. To simplify the graph, a distance threshold value was set. As a result of the work, it is revealed that the obtained characteristics indicate greater centrality and connectivity of the Nordic region graph in comparison with the Siberian region graph.*

Использование теории графов позволяет осуществить анализ данных и сравнение между регионами разных стран. Это позволяет получить более полное представление о состоянии экономики и ее взаимосвязях между различными регионами.

Сопоставление экономических показателей включает анализ различных характеристик, статистических данных и показателей экономики с целью выявления сходств, различий, тенденций и влияния

одних параметров на другие. Использование графов для визуализации экономических показателей различных регионов на одном графике является эффективным подходом.[1]

Таким образом, целью данной работы является исследование возможностей теории графов для выявления сходств или различий между социально-экономическими показателями Сибирского и Скандинавского регионов. Задача состоит в поиске характеристик графов с разными матрицами расстояний и последующим анализом полученных результатов.

Предлагается использовать относительные темпы прироста для анализа социально-экономических показателей в динамике. Это позволяет перейти к бинарным данным, где учитывается только положительный или отрицательный прирост. Рассчитывается расстояние Жаккара для сравнения этих данных. Также используются расстояния на основе коэффициента корреляции Пирсона  $r_{xy}$ , вычисляемые как  $(1 - r_{xy}^2)^{1/2}$

Для определения порога расстояния используется квантиль  $q$  заданного уровня ( $q=0.4$ ) по всем значениям из матрицы попарных расстояний. Все значения расстояний, которые больше этого порога, удаляются из матрицы смежности. Ребра графов взвешиваются на величину расстояния между соответствующими вершинами, деленного на максимальное попарное расстояние между вершинами. Ко всем значениям весов прибавляется малое значение (0.01), чтобы избежать проблем с нулевыми весами.

Данный подход иллюстрируется на примере анализа социально-экономических данных [2] мегарегиона Сибири, включающему Тюменскую область, Сибирский ФО, Дальневосточный ФО, а также Нордического региона [3], включающему в себя Данию, Финляндию, Исландию, Норвегию и Швецию. Вершины графов представляют собой показатели социально-экономической сферы, а ребра отражают их взаимосвязь.

В качестве основных характеристик, описывающих графы, выбраны следующие:

1) *Характерная длина пути (Median distance).*

Медианная длина пути, соединяющего каждую вершину с остальными.

2) *Коэффициент кластеризации (Clustering coefficient)*

Мера степени кластеризации, определяющая тесно связанные группы вершин.

3) *Центральность по собственному вектору (Eigenvector centrality)*

Мера влияния вершины в графе, основанная на соединениях с высоким весом.

4) *Центральность по близости (Closeness)*

Мера центральности вершины в сети, измеряющая ее близость ко всем остальным вершинам.

5) *Диаметр (Diameter)*

Максимально возможное расстояние между двумя вершинами графа

Рассчитанные значения характеристик представлены в таблице 1.

**Таблица 1 – Характеристики взвешенного графа**

Характеристики	Матрицы расстояний			
	Сибирского региона		Нордического региона	
	Жаккара	корреляции Пирсона	Жаккара	корреляции Пирсона
<i>Median distance</i>	0,736	0,391	0,958	1,409
<i>Clustering coefficient</i>	0,676	0,445	0,743	0,490
<i>Eigenvector centrality</i>	21,693	19,741	31,866	33,393
<i>Closeness</i>	0,025	0,043	0,014	0,009
<i>Diameter</i>	2,173	1,550	2,353	1,962

Выводы из проведенного исследования говорят о различии в связности, распределении узлов, группировке и важности узлов между графами Нордического и Сибирского регионов.

Граф Нордического региона оказывается более связным, имеет больший диаметр и содержит большее количество ребер, что указывает на наличие крепкой связи между его узлами. Также граф Нордического региона проявляет более сильную группировку узлов внутри подграфов и имеет высокую центральность по близости, что может указывать на более высокую позицию некоторых узлов в графе. Важность узлов, оцениваемая с помощью центральности по собственному вектору, также оказывается выше в графе Нордического региона.

В целом, все эти характеристики говорят о большей центральности и связности графа Нордического региона в сравнении с графом Сибирского региона. В дальнейшем полученные результаты будут учтены при проведении более глубоких исследований для экономических показателей Сибири и Скандинавии, а также их оптимальной визуализации.

Литература:

1. Timofeeva A. Sensitivity of clustering algorithms to sparseness of one correlation network / A. Timofeeva, N. Gorbacheva. – DOI 10.1007/978-3-031-36625-3\_26. – Text : direct // Lecture Notes in Computer Science. – 2023. – Vol. 13969 : Advances in swarm intelligence : proc. 14 intern. conf. ICSI–2023, China, Shenzhen, 14–18 July 2023, pt. 2. – P. 320–331

2. ЕМИСС [Электронный ресурс] / Единая межведомственная информационно-статистическая система. — Режим доступа: <https://www.fedstat.ru/>.

3. Nordic Statistics, <https://pxweb.nordicstatistics.org/pxweb/en/Nordic%20Statistics/>, last accessed 2023/02/01

**Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ,**  
**АНАЛИЗ И ОБРАБОТКА ДАННЫХ**

**ФИЛЬТРАЦИЯ ШУМОВ ПРИ РАДИОЛОКАЦИОННОМ  
ОТСЛЕЖИВАНИИ НЕСКОЛЬКИХ ОБЪЕКТОВ**

**А.А. Дарбинян, А.Р. Акопян**  
**Российско-Армянский Университет,**  
**г. Ереван, alexander.hakobyan@gmail.com**

*При отслеживании нескольких объектов могут возникнуть проблемы с отсутствием измерений координат одного из них. Если предположить, что другие объекты могут отслеживать расстояние до отсутствующего объекта, возникает проблема с точностью измерений. В настоящей работе для практической аппроксимации не только объекта, измерения которого отсутствуют, но и объектов, измерения которого известны, был использован комплекс методов, включающий расширенный фильтр Кальмана, построенное расширение фильтра UFIR и их гибридная схема. В конце работы были предоставлены конкретные примеры использования таких фильтров с соответствующими результатами и анализом точности.*

*When tracking multiple objects one can face issues with the absence of measurements of the coordinates of one of them. Assuming the other objects can track the distance from the missing one, a problem occurs regarding the accuracy of the measurements. In the current work to give a practical approximation for not only the object, which measurements are missing but also the objects which measurements are known, a complex of methods were used which include the Extended Kalman Filter, a constructed extension of UFIR filter and the hybrid scheme for those 2 filters. At the end of the work concrete examples of usages for such filters were provided with their corresponding results and precision analysis.*

Работа посвящена оценке изменяющегося  $n$ -мерного вектора состояния на основе полученных  $r$ -мерных данных от приемника. Обозначив вектор состояния в момент времени  $T_k$  как  $X_k$ , предположим, что он изменяется по следующему закону

$$X_{k+1} = F_k X_k + G_k W_k + u_k, \quad k = 0, 1, 2, \dots$$

с известным нулевым состоянием  $X_0$ , где  $F_k$  - матрица размера  $n \times n$ , называемая матрицей перехода,  $G_k$  - вещественная матрица размера  $n \times m$ ,  $u_k$  - вектор управления размера  $n$ , а  $W_k$  -  $m$ -мерный гауссовский случайный вектор с нулевым средним значением. (подробнее см. [1] и [2]). Обозначим  $Q_k$  как ковариационную матрицу  $W_k$ .

Обратите внимание, что вектор состояния может содержать информацию о нескольких движущихся объектах. В текущей статье мы

рассмотрим простой сценарий, когда вектор состояния содержит данные от трех движущихся объектов, как показано ниже:

$$X_{k+1}^{(p)} = F_k^{(p)} X_k^{(p)} + G_k^{(p)} W_k^p, \quad p = 1, 2, 3, \quad k = 0, 1, \dots$$

где векторы состояния  $X_k^{(p)}$ , матрицы  $F_k$  и  $G_k$  имеют следующий вид:

$$X_k^{(p)} = \begin{bmatrix} x_k^{(p)} \\ y_k^{(p)} \\ \dot{x}_k^{(p)} \\ \dot{y}_k^{(p)} \end{bmatrix}, \quad F_k^{(p)} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad G_k^{(p)} = \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix}, \quad X_k := \begin{bmatrix} X_k^{(1)} \\ X_k^{(2)} \\ X_k^{(3)} \end{bmatrix}.$$

Обозначим

$$F_k := \begin{bmatrix} F_k^{(1)} & 0 & 0 \\ 0 & F_k^{(2)} & 0 \\ 0 & 0 & F_k^{(3)} \end{bmatrix}, \quad G_k := \begin{bmatrix} G_k^{(1)} & 0 & 0 \\ 0 & G_k^{(2)} & 0 \\ 0 & 0 & G_k^{(3)} \end{bmatrix}, \quad p = 1, 2, 3.$$

Рассматривается следующая нелинейная модель измерений:

$$\check{X}_k = (x_k^{(1)}, y_k^{(1)}, x_k^{(2)}, y_k^{(2)}, d_k^{(1)}, d_k^{(2)})^T + V_k, \quad k = 0, 1, \dots,$$

где

$$d_k^{(p)} := \sqrt{(x_k^{(p)} - x_k^{(3)})^2 + (y_k^{(p)} - y_k^{(3)})^2}, \quad p = 1, 2,$$

а  $V_k - r$  - 6-мерный гауссовский случайный вектор с нулевым средним значением (обозначим его ковариационную матрицу как  $R_k$ ). Такое измерение представляет собой координаты двух отслеживаемых объектов и измерения расстояний этих объектов до третьего, как показано на рисунке ниже.

Используя расширенный фильтр Кальмана (РФК), нововведенный расширенный несмещенный фильтр с конечной импульсной характеристикой (РНФКИХ) и их гибридную схему удалось достигнуть существенных результатов в оптимальной фильтрации шумов как при отсутствии случайных маневров объектов, так и при наличии (см. [3]-[6]).

Данные способы фильтрации были проверены на компьютерной симуляции, результаты проверки предоставлены ниже для двух случаев – в первом случае маневры отсутствуют, во втором же случае каждый из объектов в определенный момент времени производил маневр.

Ниже можете найти результаты симуляции:



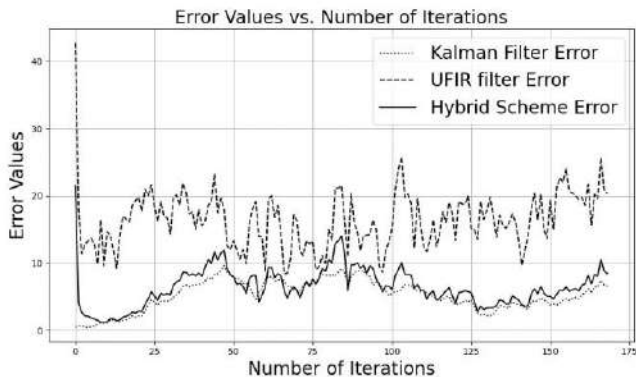


Рис. 1 – Маневры отсутствуют

Также приводятся усредненные ошибки каждого из способов фильтрации:

**Таблица 1-Ошибки при отсутствии маневров**

	РФК	РНФКИХ	Гибридная Схема
Ошибки	130.2642	21.7777	19.3217

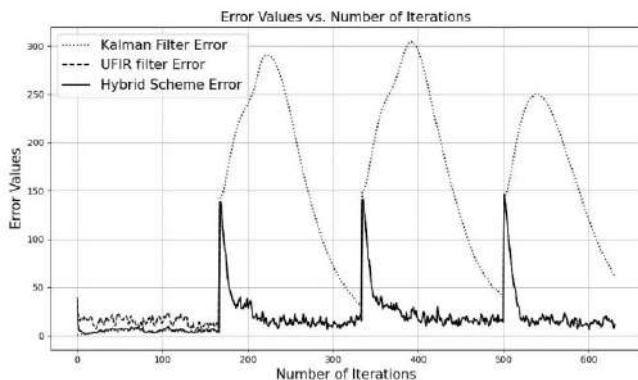


Рис. 2 – Маневры присутствуют

**Таблица 2-Ошибки при наличии маневров**

	РФК	РНФКИХ	Гибридная Схема
Ошибки	5.2262	16.4069	6.3689

Литература:

1. R. E. Kalman (1960). "A new approach to linear filtering and prediction problems". Journal of Basic Engineering. 82 (1): 35.45. doi:10.1115/1.3662552.

2. R. E. Kalman and R. S. Bucy. "New Results in Linear Filtering and Prediction Theory", 1961.
3. Y. S. Shmaliy, S. Zhao and C. K. Ahn, "Unbiased Finite Impulse Response Filtering: An Iterative Alternative to Kalman Filtering Ignoring Noise and Initial Conditions," in IEEE Control Systems Magazine, vol. 37, no. 5, pp. 70-89, Oct. 2017, doi: 10.1109/MCS.2017.2718830.
4. Unbiased FIR Filtering: An Iterative Alternative to Kalman Filtering Ignoring Noise and Initial Conditions. October 2017 IEEE Control Systems Magazine 37(5) DOI: 10.1109/MCS.2017.2718830
5. A. Darbinyan, A. R. Hakobyan and M. A. Khachaturyan, "About One Hybrid Scheme Of Unbiased Finite Impulse Response Filter And Kalman Filter". Vestnik RAU Physical, Mathematical and Natural Science № 1, pp. 5-14, 2021.
6. A. Darbinyan And A. R. Hakobyan, "Kalman Filter Modification For Polar And Spherical Coordinate Systems". Vestnik RAU Physical, Mathematical and Natural Science № 2, pp. 5-18, 2019.

## **РАЗРАБОТКА АЛГОРИТМА ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЯ ПРИ РЕНТГЕНОГРАФИИ ОБЪЕКТОВ ВЫСОКОЙ ПЛОТНОСТИ**

**А.А.Зотов**

**Новосибирский государственный технический университет,  
г. Новосибирск, kalodor@mail.ru,**

**Научный руководитель: Григорьев Д.Н., к.ф.-м.н., доц., доцент  
кафедры ЭФУиУ**

*В настоящее время ведется разработка нового детектора на основе сцинтилляционных кристаллов для проведения рентгенографических исследований объектов с высокой оптической плотностью. Для регистрации света в них используются многоэлементные КМОП-фотоприемники. Так как на одно волокно приходится примерно 4500 пикселей и положение оптических волокон относительно пикселей фотоприемника задается конструкцией крепления недостаточно точно, то необходимо разработать алгоритм для восстановления изображений, получаемых с многоэлементных фотоприемников. В данной работе представлен алгоритм, разработанный для анализа данных, описана логика его работы и полученные результаты.*

*Currently, a new detector based on scintillation crystals is being developed for X-ray studies of objects with high optical density. To register light in them, multi-element CMOS photodetectors are used. Since there are approximately*

4,500 pixels per fiber and the position of the optical fibers relative to the photodetector pixels is not set precisely enough by the mounting design, it is necessary to develop an algorithm for restoring images obtained from multi-element photodetectors. This paper presents an algorithm developed for data analysis, describes the logic of its operation and the results obtained

Рентгенография – это исследование внутренней структуры объекта путем просвечивания его рентгеновскими лучами. На Рисунке 1 представлена схема проведения рентгенографических исследований [1].

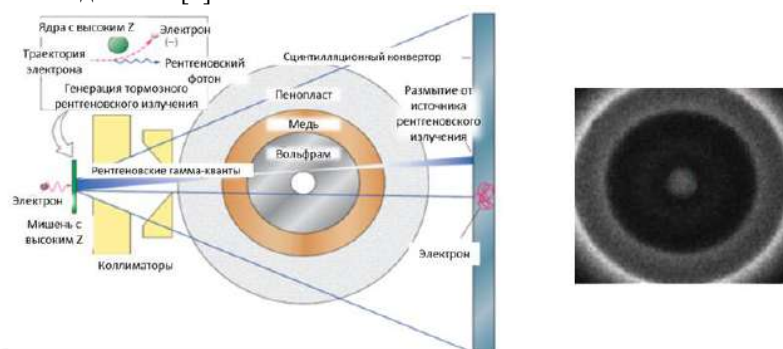


Рис. 1 – Схема проведения рентгенографических исследований(слева), рентгенограмма исследуемого объекта(справа)

Традиционно для проведения рентгенографических исследований используется метод, в котором свет со сцинтилляционного конвертора передается на матричный фотоприемник с помощью объективов. Основной недостаток такой схемы, малая величина светосбора, которая составляет не более 0.2. Замена традиционной оптической системы на волоконную оптику позволила увеличить величину светосбора в несколько десятков раз.

Разработанный детектор для проведения рентгенографических исследований состоит из сцинтилляционного конвертора, волоконной оптики и многоэлементных фотоприемников. Свет, возникший в результате взаимодействия гамма-квантов с веществом сцинтилляционного конвертора, передается по волоконной оптике к многоэлементным фотоприемникам. Положение оптических волокон относительно пикселей фотоприемника задается конструкцией крепления недостаточно точно. Количество пикселей, соответствующее одному волокну на изображении, равно 4600 штук. Поэтому была поставлена задача: разработать автоматизированный алгоритм обработки данных с фотоприемников. Созданный автоматизированный алгоритм повысил скорость обработки

информации на несколько порядков по сравнению с ручной обработкой сигналов.

Логика работы алгоритма для восстановления изображений состоит из нескольких этапов. На первом этапе из снимков, полученных с фотоприемников, сначала вычитается уровень нулевого сигнала. На втором этапе проводится цифровая обработка снимков, выделяющая границы волокна. На третьем этапе границы сохраняются и составляется карта волокон: каждому волокну присваивается уникальный индекс. Четвертый этап заключается в том, что сигналы с волокон, которые приходят на один сцинтилляционный кристалл суммируются и таким образом определяется величина сигнала в кристалле. Совокупность величин сигналов в кристаллах создает рентгенограмму исходного изображения.

Для проверки работы алгоритма и измерения пространственного разрешения макета детектора был проведен эксперимент с использованием вольфрамового сегмента. Была проведена серия измерений при перемещении сегмента в горизонтальном направлении поперёк потока с шагом 0,2 мм, как показано красной стрелкой на Рисунке 2.

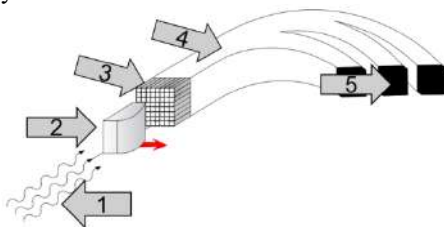


Рис. 2 – Схематическое построение эксперимента по измерению пространственного разрешения макета детектора: 1 – поток гамма-квантов; 2 – вольфрамовый сегмент; 3 – сцинтилляционный конвертор; 4 – волоконная оптика; 5 – многоэлементные фотоприемники

Полученные в результате работы алгоритма, рентгенограммы аппроксимировались функцией  $\text{erf}(x)$  (Рисунок 3).

Измеренное размытие границы сегмента составило  $1,05 \pm 0,03$  мм, что согласуется с численным моделированием физических процессов взаимодействия  $\gamma$ -квантов с веществом.

На Рисунке 4 представлена зависимость восстановленного положения границы от смещения сегмента, аппроксимированная линейной функцией. Пространственное разрешение оценивалось по распределению разностей восстановленной координаты и смещения сегмента и составило менее 0,04 мм.

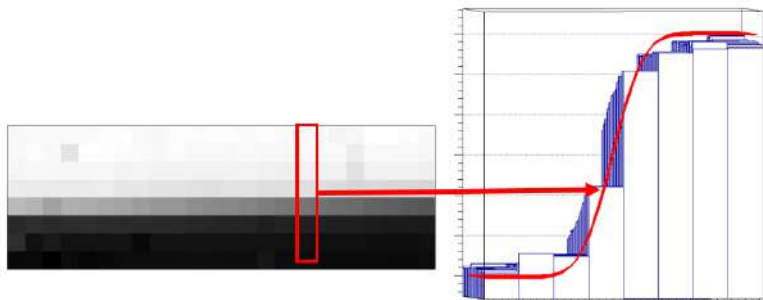


Рис. 3 – Рентгенограмма вольфрамового сегмента (слева), Аппроксимация рентгенограммы с помощью функции  $\text{erf}(x)$  (справа)

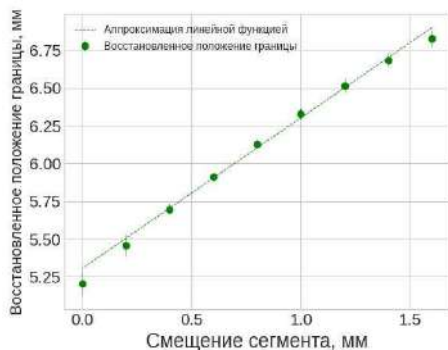


Рис. 4 – Зависимость восстановленного положения границы от смещения сегмента

В ходе работы был разработан автоматизированный алгоритм восстановления сигналов с изображений, получаемых с многоэлементных фотоприемников при рентгенографических исследованиях объектов. Данный алгоритм применен к данным, полученным в результате эксперимента по измерению пространственного разрешения макета сцинтилляционного детектора с высоким разрешением. Полученные результаты согласуются с моделированием физических процессов взаимодействия  $\gamma$ -квантов с веществом.

Литература:

1. Flash Radiography. Historical Origins. Cunningham G. S., Morris C. // Los Alamos Science. – 2003. – nr 28. – P. 76–91. – URL: <https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR->

03-5305. – Дата обращения: 25.05.2023.

2. X-ray detectors for digital radiography / M. J. Yaffe, J. A. Rowlands. – Текст : электронный // Physics in Medicine and Biology. – 1997. – Vol. 42, nr 1. – P. 1–39. – URL: <https://doi.org/10.1088/0031-9155/42/1/001>. – Дата публикации: 16.08.1997.

3. Multi-MeV Flash Radiography in Shock Physics Experiments: Specific Assemblages of Monolithic Scintillating Crystals for Use in CCD-Based Imagers / D. Partouche-Sebban, I. Abraham, S. Lauriot, C. Missault. – Текст : электронный // X-Ray Optics and Instrumentation. – 2010. – Vol. 2010. – P. 1–9. – URL: <https://doi.org/10.1155/2010/156984>. – Дата публикации: 17.01.2010.

## **ВЕБ-ТРЕНАЖЕР ДЛЯ ОБУЧЕНИЯ МЕТОДАМ ХРАНЕНИЯ ДАННЫХ В ОПЕРАТИВНОЙ ПАМЯТИ**

**В.Г. Кобылянский, И.А. Кучма**

**Новосибирский государственный технический университет,  
г. Новосибирск, [mistersandsnake@gmail.com](mailto:mistersandsnake@gmail.com)**

**Научный руководитель: Кобылянский В.Г., к.т.н., доцент**

*Отсутствие осведомленности о внутреннем устройстве типов данных является одной из проблем в знаниях современных студентов-программистов. В данной статье предлагается ее решение в виде веб-приложения для изучения представления простых и сложных типов данных в оперативной памяти компьютера.*

*The lack of knowledge about the inner working of data types is one of the problems in modern programmer students' expertise. In this article, a solution to it is provided in the form of a web application designed for learning the representation of simple and complex data types in a computer's memory.*

Типы данных являются одним из базовых понятий программирования, понимание которого является критически важным для решения большинства задач, включающих в себя низкоуровневое программирование. В большинстве случаев при изучении типов данных основное внимание уделяется семантике и синтаксису их описания, диапазону допустимых значений, набору допустимых операций и т.д. Изучению внутренних форматов хранения различных типов данных при этом уделяется незначительное внимание, что связано, на наш взгляд, с двумя причинами.

Первой причиной является широкая популярность и распространенность языков программирования с динамической типизацией, таких как Python или JavaScript, документация которых не дает почти никакой информации о внутреннем устройстве своих типов данных и информирует программиста лишь о том, как с этими типами данных работать [1-2]. Вторая причина связана с непониманием того, что любые данные обрабатываются процессором и именно он часто накладывает определенные ограничения на их внутреннюю структуру.

Это приводит к пробелу в знаниях, который усложняет работу в других языках программирования, а также решение задач, тесно связанных со знанием типов данных.

В целях решения данной проблемы было разработано веб-приложение, являющееся тренажером для обучения тому, как различные типы данных хранятся в оперативной памяти. Тренажер поддерживает большое количество простых типов данных из языков программирования C# и Pascal. Помимо того, в тренажер встроены классы C# и записи Pascal, на примере которых возможно изучение структуры сложных типов данных в этих языках.

В приложении реализованы режим обучения, позволяющий увидеть, как типы данных хранятся в памяти компьютера, и режим контроля, в котором пользователь должен определить случайно сгенерированное значение, соответствующее выбранному типу данных, по его внутреннему представлению. Введенное пользователем значение сравнивается со сгенерированным, и если они совпадают в пределах заданной точности, ответ считается верным. В ином случае, пользователю предлагается ввести другой ответ.

Для получения представления различных типов данных был разработан алгоритм, использующий битовые операции для побитового считывания значения, принадлежащего этому типу данных, в строку. Поскольку битовые операции в большинстве языков программирования возможно использовать исключительно для целочисленных типов данных, к иным типам данных применяются каламбуры типизации, позволяющие интерпретировать эти типы данных как целочисленные, не изменяя значения бит в них [3]. Такой способ применим не ко всем типам данных, встроенным в приложение, и некоторые из них потребовали использования особых алгоритмов, включающих в себя прямое считывание памяти - небезопасной операции в C#, основном языке разработки приложения. Тем не менее, использование каламбуров типизации позволило значительно сократить количество кода в приложении.

Главное окно приложения представлено на рисунке 1.

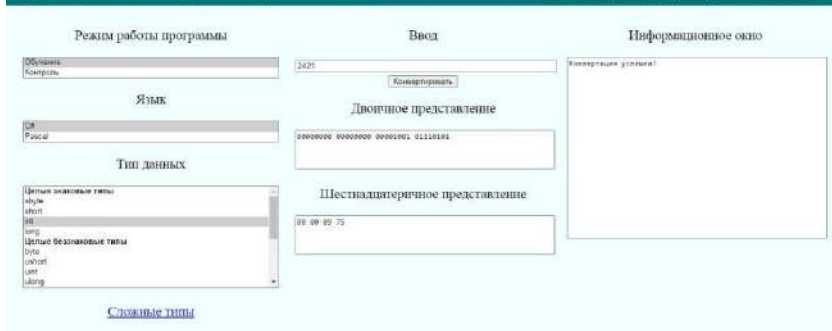


Рис. 1 – главное окно приложения

В приложении предоставляется возможность выбора режима работы программы, языка программирования и его типа данных. Пользователь может вводить любые значения, соответствующие выбранному типу. Кроме того, в приложение включены кодовые страницы для символьных типов данных и руководство пользователя, являющиеся подспорьем в изучении устройства типов данных.

В сети Интернет не было найдено ни одного аналога нашему приложению. Существуют инструменты, позволяющие изучить внутреннее представление определенных типов данных, однако информацию о других типах данных можно найти лишь глубоко в документации языков программирования, поддерживающих эти типы. Также нет иных приложений, которые позволяли бы изучать устройство сразу нескольких принципиально различающихся типов данных. Исходя из всего этого, можно судить об инновационности тренажера.

На данный момент функционал приложения полностью реализован и оно готово к использованию в учебном процессе.

#### Литература:

1. Data model [Electronic resource] // Python.org – URL: <https://docs.python.org/3/reference/datamodel.html> (accessed: 08.11.2023)
2. Типы данных [Электронный ресурс] // JavaScript.ru. Дата обновления: 26.08.2023. URL: <https://learn.javascript.ru/types> (дата обращения: 08.11.2023)
3. Кучма И.А. Теория и практические применения каламбуров типизации // Наука. Технологии. Инновации: материалы всерос. науч. конф. молодых ученых: в 11 ч. – Новосибирск: Изд-во НГТУ, 2022. – Ч. 2. – С. 27–31.



## **МОДЕЛИРОВАНИЕ ЦИФРОВОГО СЕРТИФИКАТА С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ БЛОКЧЕЙН**

**Р.А. Легалов**

**Новосибирский государственный технический университет,  
г. Новосибирск, legalov.roman@gmail.com,  
Научный руководитель: Авдеенко Т.В., д.т.н., профессор**

*В данной статье рассматривается использование технологии блокчейн для цифрового сертифицирования. Описывается принцип работы блокчейна, его преимущества и возможности в области сертификации информации.*

*This article is about using blockchain technology for digital certification. The principle of operation of the blockchain, its advantages and capabilities in the field of information certification are described.*

Цифровое сертифицирование с помощью технологии блокчейн является одним из самых перспективных направлений в области информационной безопасности и подтверждения подлинности данных. Блокчейн — это децентрализованная система, которая позволяет хранить информацию в виде цепочки блоков, каждый из которых содержит информацию о предыдущем блоке. Это делает невозможным изменение или подделку данных, так как любое изменение одного блока повлечет за собой изменение всей цепочки, что сразу же будет замечено всеми участниками сети [1].

Одним из основных преимуществ использования блокчейн для цифрового сертифицирования является отсутствие необходимости в посредниках или доверенных лицах. Это позволяет существенно уменьшить затраты на сертификацию и ускорить процесс подтверждения подлинности данных. Кроме того, блокчейн обеспечивает высокий уровень защиты информации, так как данные хранятся в распределенной сети, что делает их уязвимыми к взлому или кибератакам.

Сам же цифровой сертификат — это электронный документ, который подтверждает подлинность и достоверность информации о чем-либо, например, о человеке, компании или товаре. Он хранится в блокчейн-сети, которая является децентрализованной и надежной системой хранения данных [2].

Процесс работы цифрового сертификата начинается с создания уникальной цифровой подписи, которая связывается с определенным объектом или субъектом. Эта подпись затем записывается в блокчейн-реестр, где она становится неизменяемой и доступной для проверки всем участникам сети.

Когда кто-то хочет проверить подлинность цифрового сертификата, он может обратиться к реестру и убедиться, что цифровая подпись соответствует исходным данным. Таким образом, блокчейн обеспечивает прозрачность и надежность цифровых сертификатов, предотвращая возможные фальсификации или изменения информации.

Для обеспечения безопасности самого блокчейна используется хэширование. Хэширование — это процесс преобразования данных в непредсказуемую строку фиксированной длины, которая называется хеш-значением. Это делается с помощью хэш-функций, которые принимают входные данные и возвращают уникальное хеш-значение. В блокчейне хэширование используется для обеспечения целостности данных и защиты от подделки. Каждый блок в блокчейне содержит хеш предыдущего блока, что создает цепочку блоков и делает невозможным изменение данных в прошлых блоках без изменения всех последующих блоков. Примерами хэш-функций, применяемых в блокчейне, являются SHA-256, SHA-3, RIPEMD-160 и другие. Эти функции обладают свойством высокой степени уникальности и стойкости к коллизиям, что делает их подходящими для использования в блокчейне.

В данной работе рассмотрены несколько примеров применения цифровой сертификации, а также алгоритмы работы с цифровыми сертификатами на основе блокчейн технологии:

1. Цифровое сертифицирование полной цепочки поставок: от поставщика материалов до конечного покупателя.

Отслеживание контрафактных товаров стало более эффективным благодаря использованию цифровых сертификатов, основанных на блокчейн технологии [3]. Благодаря этому инновационному подходу, каждый этап производства и поставки товара может быть зафиксирован в цепи блоков, что обеспечивает прозрачность и надежность данных. В любой момент поставки определенного товара можно удостовериться в его оригинальности, просканировав цифровой сертификат с помощью мобильного приложения или специального устройства [4]. Это позволяет бороться с контрафактом, защищать права потребителей и повышать доверие к брендам и производителям. К примеру, поставщик сырья (материалов) может отследить, сколько сделано и продано товара определенным производителем, и тем самым выяснить, не покупает ли производитель серое (незадекларированное или же плохого качества) сырьё, чтобы изготавливать больше продукции по более низкой цене. Производитель же сможет узнать о том, что магазин продаёт его товар в том количестве, в котором поставляет завод, а не более, продавая часть поддельных товаров путем смешивания их с оригинальным. Покупатель, в свою очередь сможет проследить полный путь товара, от

завода до магазина, и тем самым подтвердить, что это действительно оригинальный товар, а не контрафакт. Вся цепочку взаимодействий между контрагентами и покупателем можно просмотреть на рисунке 1.

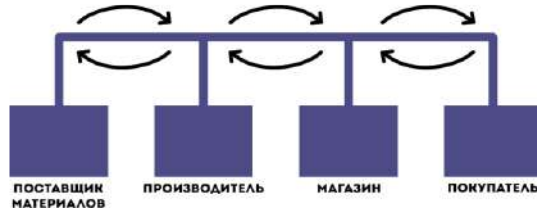


Рис. 1 – Цепочка поставки оригинального товара от поставщика материалов до конечного покупателя

2. Цифровое сертифицирование полной цепочки поставок редких вещей от владельца до конечного покупателя.

В мире достаточно много вещей, существующих в единственном или в очень малом количестве экземпляров. Для того, чтобы подтвердить подлинность таких вещей также можно использовать цифровые сертификаты на основе блокчейн технологии, так как их нельзя подделать [5]. Для владельца вещи цифровой сертификат даёт доказательство её подлинности. Для того, чтобы получить сертификат, вещь отправляется на оценку независимому эксперту, который подтверждает или отрицает подлинность. Для минимизации рисков можно добавить в цепочку ещё несколько экспертов, которые никак друг с другом не связаны (некая децентрализация в реальной жизни). После получения сертификата владелец может продать товар покупателю, и покупатель сможет проверить его подлинность. Это поможет не только защитить покупателя от подделки, но и не оставит возможности владельцу продать товар несколько раз (копию товара, затем оригинал или же вообще набор копий). Цифровой сертификат будет действителен только на один товар, и в нём будет запись о том, что позиция уже продана. Данный алгоритм можно рассмотреть на рисунке 2

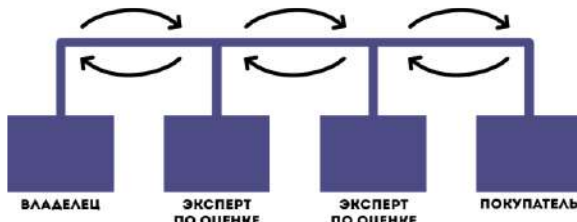


Рис. 2 – Алгоритм сертификации редких вещей

Оба примера, если они будут разработаны с использованием цифрового сертификата и блокчейна, будут работать при следующих условиях:

1. Полная децентрализация самого блокчейна. Это гарантирует невозможность редактирования общего реестра.
2. Запись всех действий, связанных с предметом, в блокчейн, конкретнее, любое считывание сертификатов (в момент покупки, продажи, перевозки, экспертизы и других действий).

Саму же модель поставки можно выразить формулой 1

$$(\forall t \in T)(\exists \text{ количество}(t) = \epsilon) \rightarrow (\text{поставка})(\forall t \in T)((\exists \text{ количество}(t) = (\epsilon' \cup \epsilon) \cap (\epsilon \neq \epsilon')) \quad (1)$$

В данной формуле символ  $\epsilon$  обозначает количество товара  $T$  ( $\epsilon'$  – другое количество). После транспортировки и складирования товара – его количество может отличаться. Это обозначает, что товар разделен на партии, либо использован частично. Цифровой сертификат вычисляет эти количества и обновляет данные в блокчейне. Когда поставка окончена цифровой сертификат подсчитывает весь товар, и в этом случае сумма его должна равняться  $\epsilon$ .

В заключение, блокчейн представляет собой мощный инструмент для цифрового сертифицирования, обеспечивая высокий уровень защиты данных и отсутствие необходимости в посредниках. Математическое моделирование и анализ данных могут помочь обеспечить надежность и безопасность блокчейна, предсказывая возможные уязвимости и атаки, а также оптимизируя алгоритмы и протоколы. Например, с помощью математических моделей можно провести анализ вероятности успешной атаки на блокчейн, определить оптимальные параметры для защиты сети и разработать новые методы шифрования данных. Также с их помощью можно проводить симуляции различных сценариев работы блокчейна, что помогает выявить потенциальные проблемы и недочеты в системе. В целом, анализ данных и математическое моделирование позволяет создать более надежные и безопасные блокчейн-системы, что является критически важным для защиты цифровых активов и данных пользователей. В будущем, блокчейн может стать основой для цифрового сертифицирования во многих сферах, от финансовых операций до медицинской документации, обеспечивая прозрачность, надежность и подлинность информации.

Литература:

1. Laurence T. Blockchain for Dummies. – United States, 2017. – 240 p.
2. Шамсутдинова Т.М. Применение технологии блокчейн для выдачи цифровых дипломов: проблемы и перспективы [Электронный ресурс] // Открытое образование: электрон. науч.-техн. журн. – № 12. – 2012. – URL: <https://cyberleninka.ru/article/n/primeneniye-tehnologii-blokcheyn-dlya-vydachi-tsifrovyyh-diplomov-problemy-i-perspektivy> (дата обращения: 25.10.2023). – doi: 10.21686/1818-4243-2018-6-51-58.
2. Laurence T. Blockchain for Dummies. – United States, 2017. – 240 p.
3. Генкин А.С., Михеев А.А. Блокчейн. Как это работает и что ждет нас завтра. – М.: Альпина Паблишер, 2018. – 592 с.
4. Кудряшов С.Ю. Применение технологии блокчейн с целью повышения доверия пользователей к использованию юридически значимых баз данных: дис. ... канд. техн. наук. – Москва, 2021. – 82 с.
5. Лелу Л. Блокчейн от А до Я. – М.: Бомбора, 2018. – 256 с.
6. The Blockchain-Based Digital Certificate for the Transport of Dangerous Goods / Imeri A., Feltus C., Agoulmine N., Khadraoui D. // Blockchain Driven Supply Chains and Enterprise Information Systems (c). – 2022. – P. 43-61. – doi: 10.1007/978-3-030-96154-1\_3

## **ИССЛЕДОВАНИЕ КОМПЛЕКСНЫХ ГЕННЫХ СЕТЕЙ: БАЗОВЫЕ МЕТОДИКИ И ИХ ПРИМЕНЕНИЕ**

**А.М. Мироненко**

**Новосибирский государственный технический университет,**

**г. Новосибирск, a\_lisa\_mironenko@mail.ru**

**Научный руководитель: Карманов В.С., к.т.н.**

*Разрабатывается математическое, алгоритмическое и программное обеспечение для моделирования сложных генных сетей с помощью метода мозаичного моделирования. Этот метод позволяет определить поведение другой части или всей сети, исходя из фрагмента генной сети. Для решения задачи моделирования сложных сетей используется подход, основанный на базовых формах. В статье представлены примеры моделирования, основанные на трех простых базовых моделях.*

*Mathematical, algorithmic and software are being developed for modeling complex gene net-works using the mosaic modeling method. This method allows you to determine the behavior of another part or the entire network based on a fragment of the gene network. To solve the problem of modeling complex networks, an approach based on basic forms is used. The article presents modeling examples based on three simple basic models.*

## Введение

Разработка математических моделей в биологических системах – актуальное направление в естественных науках. Изучается регуляция жизнедеятельности клеток через взаимодействие генов в генных сетях. Генные сети формируют фенотипические характеристики организмов на основе геномной информации.

В данной работе используется подход мозаичного моделирования, основанный на моделировании участка генной сети с учетом взаимодействий между генами и принципов их регуляции [1, 2]. Регуляция генов направлена на поддержание стационарного уровня экспрессии генов (процесс, в ходе которого наследственная информация от гена преобразуется в функциональный продукт - РНК или белок) и требует баланса взаимодействий между ними [3, 4].

При использовании подхода мозаичного моделирования возможно решение следующих задач:

1. Моделирование генной сети по известной структуре модели и её параметрах, определение динамики фазовых переменных, нахождение стационарных режимов.
2. Построение функции взаимодействия по наблюдению какой-либо известной подсети для прогнозирования состояния всей сети.
3. На основе известных закономерностей определение степени воздействия на элементы сети и изменение её параметров.

Цель данной работы состоит в разработке математического и алгоритмического решения задачи моделирования генных подсетей на основе базовых форм, на примере сети из трех вершин. Данная работа является продолжением исследования [5].

## Материалы и методы

Генная сеть (ГС) представляется в виде ориентированного графа, в котором каждая вершина соответствует определенному гену, каждое ребро соответствует связи между генами. Связи могут быть активирующего типа (один ген усиливает экспрессию другого гена) и подавляющего типа (один ген замедляет экспрессию другого гена).

Математическая модель динамики генной сети описывается системой дифференциальных уравнений (1), при этом в правую часть входят обобщенные функции Хилла [1]:

$$\frac{d}{dt} X_i(t) = \mu_i \frac{1 + \sum_{j \in V_i} \alpha_{ij} X_j^{\gamma_{ij}}(t)}{1 + \sum_{j \in W_i} \alpha_{ij} X_j^{\gamma_{ij}}(t)} - \beta_i X_i(t), \quad i \in \{1..N\}, \quad (1)$$

где  $N$  – количество вершин графа, задающего топологию ГС;  
 $X_i(t)$  – концентрация продукта в вершине с номером  $i$  в момент времени  $t$ ,

$\mu_i, \beta_i, \alpha_{ij}, \gamma_{ij}$  – числовые параметры модели.

В случае решения обобщенной задачи и безразмерных (нормированных) показателей, предлагается использовать следующие допустимые диапазоны изменения параметров:

$$0 < \mu_i \leq 1, \quad 0 < \beta_i, \alpha_{ij} \leq 2, \quad \gamma_{ij} \in \{1, 2, 3\}, \quad \forall i \in \{1..N\}, \forall j \in V_i, W_i. \quad (2)$$

Параметры модели  $\gamma_{ij}$  зависят от количества стадий в сложных многоступенчатых химических реакциях, поэтому могут быть натуральными числами; параметр  $\beta_i$  соответствует коэффициенту деградации вещества в вершине графа  $i$ ; параметры  $\alpha_{ij}$  обозначают степень влияния активирующей или подавляющей связи вершины; коэффициент  $\mu_i$  является общим нормирующим множителем.

Рассмотрим простейшую базовую форму ГС, состоящую из двух вершин. Граф сети показан на Рис. 1.



Рис 1. Граф базовой формы геной подсети, состоящий из а) двух активирующих вершин; б) активирующей и подавляющей вершины; в) двух подавляющих вершин.

Математическая модель динамики ГС заданной топологии для двух активирующих вершин представляется системой дифференциальных уравнений (СДУ) вида

$$\begin{cases} \frac{d}{dt} X_1(t) = \mu_1 \left( 1 + \alpha_{12} X_2^{\gamma_{12}}(t) \right) - \beta_1 X_1(t), \\ \frac{d}{dt} X_2(t) = \mu_2 \left( 1 + \alpha_{21} X_1^{\gamma_{21}}(t) \right) - \beta_2 X_2(t). \end{cases} \quad (3)$$

Для двух подавляющих вершин СДУ выглядит имеет следующий вид

$$\begin{cases} \frac{d}{dt} X_1(t) = \frac{\mu_1}{1 + \alpha_{12} X_2^{\gamma_{12}}(t)} - \beta_1 X_1(t), \\ \frac{d}{dt} X_2(t) = \frac{\mu_2}{1 + \alpha_{21} X_1^{\gamma_{21}}(t)} - \beta_2 X_2(t). \end{cases} \quad (4)$$

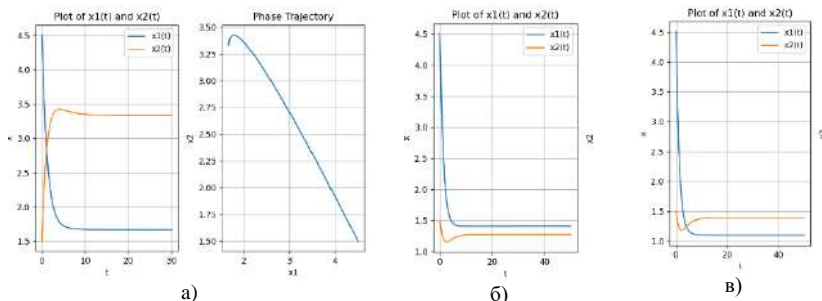
### Результаты моделирования

Рассмотрим несколько вариантов модели (3) с различными значениями параметров. Числовые значения параметров приведены в таблице 1.

**Таблица 1 – Значения параметров базовой формы модели ГС**

Вариант модели	Нач. условия	$\gamma_{12}$	$\gamma_{21}$	$\mu_1$	$\mu_2$	$\beta_1$	$\beta_2$	$\alpha_{12}$	$\alpha_{21}$
Модель 1	[4.5, 1.5]	1	1	1	1	0.8	0.5	0.1	0.4
Модель 2	[12.1, 14.9]	1	1	1	1	0.008	0.005	0.1	0.07
Модель 3	[0, 0]	2	2	1	1	0.6	0.05	0.1	0.7

Для нахождения стационарных режимов были построены соответствующие графики  $x_1(t)$  и  $x_2(t)$ , время  $t$  было выбрано исходя из физических условий задачи.



**Рис 2. – Стационарные режимы для Модели 1**

а) две активирующие вершины и график фазовой траектории; б) активирующая и подавляющая вершины; в) две подавляющие вершины

**Таблица 2 – Стационарные режимы для выбранных вариантов моделей**

Вариант модели	$X_{S_1}$	$X_{S_2}$
<i>Модель 1</i>		
Две активирующие вершины	1.66666667	3.33333333
Активирующая и подавляющая	1.40985257	1.2788206
Две подавляющие вершины	1.09746705	1.38986548
<i>Модель 2</i>	Не существует	Не существует
<i>Модель 3</i>	Не существует	Не существует

Выполнялся анализ стационарных режимов [6, 7]. Решение систем ДУ выполнялось численно с помощью метода Рунге-Кутты 4-го порядка [8]. Из таблицы 2 видно, что при наличии двух активирующих вершин выработка белкового продукта максимальна, а при двух подавляющих вершинах – минимальна, что не противоречит физике процесса.

**Заключение**

В результате выполненных исследований было разработано математическое и алгоритмическое решение задачи моделирования



генных сетей на основе базовых форм. Показано применение разработанных методов, рассмотрены три возможных случая связи двух вершин.

Полученные результаты послужат основой для перехода к следующему этапу исследований – мозаичному моделированию подсетей, управлению динамикой генных сетей, статистическому анализу параметров ГС. В дальнейшем планируется исследовать влияние каждой из компонент СДУ на выход системы на стационарный режим с помощью построения функции чувствительности для каждой из компонент.

#### Литература:

1. Olga V. Popik, Evgeny D. Petrovskiy, Elena L. Mishchenko, Inna N. Lavrik, Vladimir A. Ivanisenko Mosaic gene network modelling identified new regulatory mechanisms in HCV infection // ELSEVIER. - 2016. - №218. - С. 71-78.

2. Е.Л. Мищенко, Н.В. Иванисенко, И.Р. Акбердин, П.С. Деменков, В.А. Лихошвай, Н.А. Колчанов, В.А. Иванисенко. Подавление репликации субгеномного рнк репликаона вируса гепатита с ингибитором ns3 протеазы sch 503034 в huh-7 клетках: стохастическая модель // Вавиловский журнал генетики и селекции. - 2012. - №16. - С. 339-346.

3. О.-Я.Л.Бекиш. Медицинская биология. — Минск: Ураджай, 2000. — С. 110—111. — 518 с.

4. Цыган В. Н. Роль апоптоза в регуляции иммунного ответа // Обзоры по клинической фармакологии и лекарственной терапии. - 2004. - №2. - С. 62-78.

5. Карманов В. С. Методы моделирования сложных генных сетей на основе базовых форм = Methods of modeling complex gene networks based on basic forms / В. С. Карманов, М. Е. Царев, С. С. Хлопина. - // Обработка информации и математическое моделирование: материалы Рос. науч.-техн. конф., Новосибирск, 20–21 апр. 2022 г. – Новосибирск: Изд-во СибГУТИ, 2022. – С. 92–99. - ISBN 978-5-91434-070-1.

6. Нефедов Н. Н., Попов В. Ю., Волков В. Т. Обыкновенные дифференциальные уравнения. Курс лекций — М.: Физический факультет МГУ им. М.В. Ломоносова, 2016. — 200 с.

7. Щуров И. Обыкновенные дифференциальные уравнения Интерактивный учебник // URL: <https://ode.mathbook.info/chapter/label/chap:10prim:linearization/> (дата обращения: 01.05.2022).

8. Крайнов А. Ю., Моисеева А. Е. Численные методы решения краевых задач для обыкновенных дифференциальных уравнений. - Томск: SST, 2016. - 44 с.

## **ПРОГРАММА ДЛЯ ОБРАБОТКИ, АНАЛИЗА И ОЧИСТКИ ДАННЫХ В РАЗЛИЧНЫХ ФОРМАТАХ**

**Д.В. Пекишев, А.В. Коваленко**

**Кубанский государственный университет,**

**г. Краснодар, [anaable1974@gmail.com](mailto:anaable1974@gmail.com)**

**Научный руководитель: Ургенов М.Х., д. физ.-мат. н., профессор**

*В данной статье представлено программное средство, спроектированное для анализа, преобразования и очистки данных из различных источников, таких как файлы в форматах CSV, JSON и SQL. Программа предоставляет широкий набор функциональных возможностей, включая определение формата и типов записей, классификацию данных, а также удаление ненужных и ошибочных записей. Описаны методы работы программы и области ее применения в анализе и обработке больших объемов данных.*

*This article presents a software tool designed to analyze, transform and clean data from various sources, such as files in CSV, JSON and SQL formats. The program provides a wide range of functionality, including defining the format and types of records, classifying data, and removing unnecessary and erroneous records. The methods of operation of the program and the areas of its application in the analysis and processing of large volumes of data are described.*

Объектом исследования в данной работе являются утечки данных пользователей и способы защиты от их последствий.

Проблема состоит в большом количестве утечек личных данных, которые могут создать неприятности пользователям, если от них не защититься вовремя.

Цель работы: создание программы для анализа и обработки данных, проверки безопасности аккаунта и для устранения угроз взлома, вызванных утечками данных.

Методологическая основа исследования включает в себя статистический анализ, сравнительный анализ, экстраполяцию, моделирование, синтез выводов, анализ.

В результате работы была разработана программа, которая реализует функционал по решению задачи анализа угроз для предотвращения взлома аккаунтов.

Научная новизна работы заключается в использовании новых алгоритмов быстрого поиска по данным, а также в разработке универсальных алгоритмов унификации и разметки данных.

В результате решения поставленных задач был определен универсальный формат данных для хранения информации и поиска по ней.

Кроме того, было разработано приложение в telegram, позволяющее осуществлять мгновенный поиск по этим данным и получать результаты поиска в удобном формате.

Рассмотрим функционал программного продукта. Программа начинает свою работу с анализа входных файлов. Она определяет их формат (CSV, JSON, SQL) и начинает анализировать структуру данных. Для этого используются различные алгоритмы и методы, позволяющие выявить типы записей, их структуру и связи между ними. Далее производится очистка данных – важный этап обработки информации [1].

Программа проводит анализ каждой записи и выявляет ошибки, битые или пустые значения, дубликаты и пустые значения вроде NULL или \N. Эти записи затем удаляются или исправляются, чтобы получить чистые и надежные данные. После очистки и классификации данных программа предоставляет возможность их преобразования. При этом производится сортировка, удаление лишних столбцов и сохранение программы в унифицированном формате CSV.

Программное средство, описанное выше, имеет широкие области применения в анализе и обработке данных. В данной работе оно используется для анализа и обработки утечек данных. Утечки данных представляют серьезную угрозу для компаний и организаций, и быстрая реакция на инциденты утечки данных является важным компонентом обеспечения безопасности информации.

При обнаружении утечки данных информация может поступать из разных источников и иметь разные форматы. Программа позволяет унифицировать данные и приводить их к единообразному формату, что упрощает дальнейший анализ.

В целом использование данной программы в процессе анализа и обработки утечек данных помогает организациям быстро и эффективно реагировать на инциденты и минимизировать потенциальные угрозы для безопасности информации.[2]. Унификация, классификация и очистка данных снижают риски и обеспечивают более точный и своевременный анализ утечек данных, что важно для защиты конфиденциальной информации и поддержания доверия клиентов и партнеров

Литература:

1. Provost, F., & Fawcett, T. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media, 2013. – 413 с.

2. Inmon, W. H., & Nesavich, T. Tapping into Unstructured Data: Integrating Unstructured Data and Textual, 2008. – 241 с.

## **СРАВНИТЕЛЬНЫЙ АНАЛИЗ СКОРОСТИ ВЫЧИСЛЕНИЙ АЛГОРИТМОВ КОМБИНАТОРНОЙ ГЕНЕРАЦИИ ДЛЯ МНОЖЕСТВ, ОПРЕДЕЛЯЕМЫХ ЧИСЛАМИ ФУБИНИ**

**В.А. Полюга, Ю.В. Шабля**

**Томский государственный университет систем  
управления и радиоэлектроники,  
г. Томск, [syv@fb.tusur.ru](mailto:syv@fb.tusur.ru)**

**Научный руководитель: Шабля Ю.В., к.т.н., с.н.с. ЛАТИДС**

*В данной работе проводится сравнение скорости вычислений трех разработанных алгоритмов комбинаторной генерации для множеств, функция мощности которых определяется числами Фубини. Результаты проведенного экспериментального исследования подтвердили полученные теоретические оценки вычислительной сложности разработанных алгоритмов.*

*This paper compares the computational speed of three developed combinatorial generation algorithms for sets whose cardinality function is determined by the Fubini numbers. The results of the experimental study confirmed the obtained theoretical estimates of the computational complexity for the developed algorithms.*

Комбинаторная генерация – научное направление, изучающее комбинаторные множества и алгоритмы генерации элементов таких множеств [1-2]. Примерами комбинаторных множеств являются множества сочетаний элементов, перестановок элементов и т.д. Существует множество различных методов построения алгоритмов комбинаторной генерации и каждый из них имеет свои ограничения. Одним из таких методов является метод с использованием структур деревьев И/ИЛИ [3].

На основе метода с использованием структур деревьев И/ИЛИ в рамках предыдущих исследований были получены алгоритмы ранжирования (Rank) и генерации по рангу (Unrank) для комбинаторных множеств, функция мощности которых определяется числами Фубини, с использованием трех различных выражений

функции мощности [4-5]. Ранжирование – процесс присваивания каждому элементу комбинаторного множества индивидуального номера. Генерация по рангу – обратный процесс восстановления элемента комбинаторного множества по его номеру. Целью данной работы является сравнение скорости вычислений трех разработанных алгоритмов комбинаторной генерации, а также сопоставление результатов экспериментального исследования времени работы программной реализации алгоритмов с полученными теоретическими оценками их вычислительной сложности.

Числа Фубини, обозначим через  $F_n$ , формируют числовую последовательность, начиная с  $F_0 = 1$ . В таблице 1 представлены несколько первых значений чисел Фубини [6].

**Таблица 1 – Числа Фубини**

$n$	0	1	2	3	4	5	6	7	8	...
$F_n$	1	1	3	13	75	541	4683	47293	545835	...

Числа Фубини имеют самые различные комбинаторные интерпретации. Например, значения числа Фубини  $F_n$  можно интерпретировать как значение количества способов распределения призовых мест между  $n$  участниками с учетом ничьих.

При разработке соответствующих алгоритмов комбинаторной генерации были выбраны следующие три способа подсчета значений чисел Фубини:

1) используя формулу суммирования, включающую числа Стирлинга второго рода:

$$F_n = \sum_{k=1}^n k! S_n^k; \quad (1)$$

2) используя формулу суммирования, включающую числа Эйлера первого рода:

$$F_n = \sum_{k=0}^{n-1} 2^k E_n^k; \quad (2)$$

3) используя рекуррентную формулу, включающую биномиальные коэффициенты:

$$F_n = \sum_{k=1}^n C_n^k F_{n-k}, \quad F_0 = 1. \quad (3)$$

Все вышеописанные формулы удовлетворяют условиям выбранного метода построения алгоритмов комбинаторной генерации, а именно не противоречат алгебре  $\{N, +, \times, R\}$  (использование целых чисел, операций сложения, операций умножения, рекурсивных вызовов). Для разработанных алгоритмов комбинаторной генерации были получены

теоретические оценки вычислительной сложности с учетом их использования только для комбинаторных объектов с максимальным значением ранга при заданном значении параметра  $n$  (таблица 2).

**Таблица 2 – Вычислительная сложность алгоритмов**

	Алгоритм Rank	Алгоритм Unrank
Алгоритмы на основе формулы (1)	$O(n^3)$	$O(n^3)$
Алгоритмы на основе формулы (2)	$O(n^3)$	$O(n^3)$
Алгоритмы на основе формулы (3)	$O(n^4)$	$O(n^4)$

Для подтверждения полученных теоретических оценок вычислительной сложности был проведен вычислительный эксперимент. Все разработанные алгоритмы комбинаторной генерации были реализованы на языке программирования Python 3.10.12. В ходе проведения эксперимента определяется зависимость времени выполнения алгоритмов от значения параметра  $n$  при фиксированном максимально возможном значении ранга. Для каждого значения  $n$  в интервале от 1 до 100 с шагом 1 проводились вычисления алгоритма *Unrank* и *Rank* с повторением 500 раз. В результате бралось среднее значение времени выполнения вычислений внутри алгоритмов. На основе полученных данных были построены линии регрессии в виде полинома, порядок которого соответствует показателю степени для  $n$  в теоретической оценке вычислительной сложности алгоритма. Далее проводилось контрольное тестовое измерение времени выполнения алгоритмов для  $n = 200$ , которое сравнивалось с теоретическим прогнозом на основе поведения линии регрессии.

В качестве примера на рисунке 1 изображены результаты вычислительного эксперимента для алгоритмов на основе формулы (1). Визуальное сравнение тестового результата с линией регрессии в виде полинома 3-ей степени подтверждает полученную теоретическую оценку вычислительной сложности алгоритмов  $O(n^3)$ . Аналогичные результаты были получены для алгоритмов на основе формулы (2) и алгоритмов на основе формулы (3).

В таблице 3 приведены характеристики оборудования, на котором проводился вычислительный эксперимент.

**Таблица 3 – Характеристики оборудования**

Модель оборудования	Acer Aspire E5-576G
Оперативная память	8 Гб
Процессор	Intel Core i5-7200U CPU 2.5GHz x4
Операционная система	Ubuntu 22.04.3 LTS

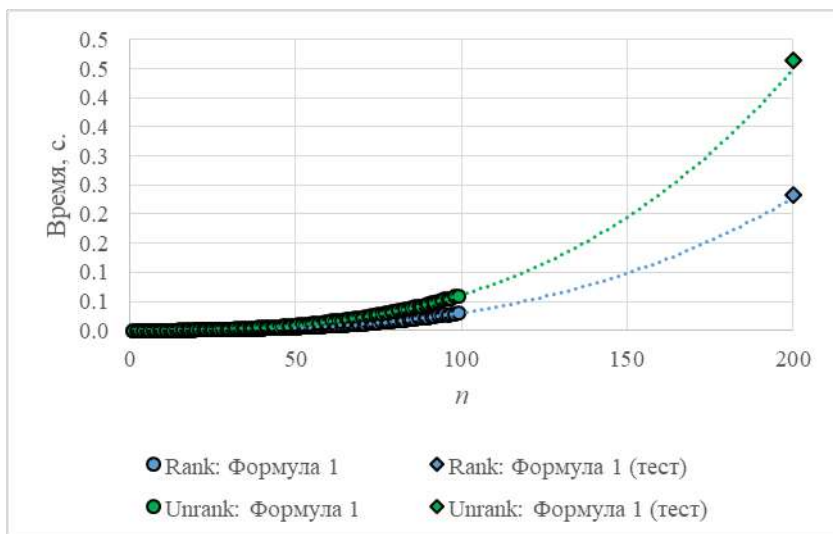


Рис. 1 – Время выполнения алгоритма на основе формулы (1)

Таким образом, менее трудоемкими с точки зрения времени вычисления оказались алгоритмы комбинаторной генерации, основанные на формулах без рекурсивных вызовов. Это вызвано сложностью вычисления чисел Фубини, требуемых при работе алгоритмов на основе формулы (3).

Работа выполнена при финансовой поддержке РФФ в рамках научного проекта № 22-71-10052.

#### Литература:

1. Кнут Д.Э. Искусство программирования. Том 4А. Комбинаторные алгоритмы, часть 1. – М.: Вильямс, 2013. – 960 с.
2. Kreher D.L. Stinson D.R. Combinatorial algorithms: Generation, enumeration, and search. – USA: CRC Press, 1999. – 329 с.
3. Кручинин В.В. Методы построения алгоритмов генерации и нумерации комбинаторных объектов на основе деревьев И/ИЛИ. – Томск: В-Спектр, 2007. – 199 с.
4. Shablya Y., Polyuga V. Development of combinatorial generation algorithms for discrete structures associated with the Fubini numbers // Proceedings book of the 13th Symposium on Generating Functions of Special Numbers and Polynomials and their Applications (11-13 March 2023). – Antalya, 2023. – P. 293-296.
5. Полюга В.А., Шабля Ю.В. Разработка алгоритмов комбинаторной генерации для множеств, определяемых числами Фубини //

Перспективы развития фундаментальных наук: сборник научных трудов XX Международной конференции студентов, аспирантов и молодых ученых (25-28 апреля 2023 г., Томск): в 7 томах. – Томск: Изд-во ТПУ, 2023. – Т. 3. Математика. – С. 37-39.

6. Fubini numbers [Электронный ресурс]. – Режим доступа: <https://oeis.org/A000670>, свободный (дата обращения: 01.11.2023).

## **ПРИМЕНЕНИЕ НЕЧЕТКОГО ВЫВОДА ДЛЯ ОЦЕНКИ ЭЛЕМЕНТОВ ГИБРИДНОЙ КОГНИТИВНОЙ КАРТЫ**

**А.А. Прокудин**

**Томский государственный университет систем  
управления и радиоэлектроники,  
г. Томск, igor.petrovich.rabinovitz@gmail.com**

**Научный руководитель: Силич М.П., д.т.н., профессор каф. АОИ**

*Данная статья рассматривает возможности применения систем нечеткого вывода для оценки на основе данных состояния факторов и силы влияния связей в гибридных когнитивных моделях. На примере показано, что использование нечеткого вывода помогает улучшить точность прогнозирования моделей, а также лучше учесть индивидуальные характеристики анализируемых объектов.*

*This article considers the possibilities of using fuzzy inference systems to assess the state of factors and their strength of influence in hybrid cognitive models on the basis of data. The provided example shows that the use of fuzzy inference helps to improve the accuracy of forecasting in models, it also makes it easier to take specific properties of the analyzed objects into account.*

Когнитивные карты являются распространенным методом моделирования слабоструктурированных социально-экономических систем. Они позволяют доступно представить существующую проблемную ситуацию в виде графа, вершинами которого являются факторы ситуации, а ребрами – предполагаемые связи между ними. При этом для оценки факторов и их связей могут использоваться нечеткие величины, что позволяет учесть существующую в сложных системах неопределенность.

В данной работе рассматривается методология иерархических гибридных когнитивных карт (ИГКК), которая позволяет проводить оценку с помощью различных способов анализа статистических данных, что должно увеличить точность оценки и сделать принятие решений на основе когнитивной модели более обоснованным [1].



Факторы модели для каждого объекта связываются с одним или несколькими индикаторами, его реальными количественными показателями. На основе всего релевантного набора данных получается нечеткая оценка для всех факторов каждого конкретного объекта (используются алгоритмы построения функций принадлежности и кластеризации). Для получения оценок силы влияния могут быть использованы регрессионные модели. Данный вид карт может использоваться для сравнительного анализа, диагностики проблем конкретного объекта и прогнозирования его будущего состояния.

ИГКК позволяют получить более объективные модели, чем подходы, основанные на экспертной оценке, но все еще остаются проблемы. Во-первых, прогнозирование с помощью линейных регрессионных моделей все еще часто дает недостаточно точные результаты. Во-вторых, сами силы влияния обычно представляются четкими оценками, хотя можно предположить, что принадлежность объекта к той или иной группе влияет на взаимоотношения факторов.

Вторая проблема решается в нечетких продукционных картах, которые для оценки силы влияния и прогнозирования используют базы правил [2]. Сами нечеткие оценки тоже часто получаются экспертным путем, но данный процесс можно автоматизировать и заменить оценки, основанные на регрессионных коэффициентах, на оценки, выводимые с помощью системы нечеткого вывода типа Такаги-Сугено. Это решает и первую проблему. Целью работы является исследование эффективности применения систем нечеткого вывода, построенных на основе данных, для задачи оценки факторов и связей гибридной когнитивной карты.

Рассмотрим следующий пример: на рисунке 1 показано несколько связанных факторов простой когнитивной карты. Для оценки используется набор статистических данных по 10 регионам Сибирского федерального округа за период с 2010 по 2021 годы. Данные взяты с сайта Федеральной службы государственной статистики [3]. Поставлена задача предсказания уровня бедности по регионам за 2021 год на основе оценок входных факторов. В таблице 1 представлены условные обозначения, используемые далее в тексте.

**Таблица 1 – Условные обозначения**

Фактор	Обозначение фактора	Обозначение индикатора
Бедность	$F_1$	$y$
Безработица	$F_2$	$x_1$
Доход населения	$F_3$	$x_2$

Данные были очищены от выбросов и нормализованы. Сначала была получена модель множественной линейной регрессии (1).

$$y(x_1, x_2) = 0.28x_1 - 0.39x_2 + 0.5 \quad (1)$$

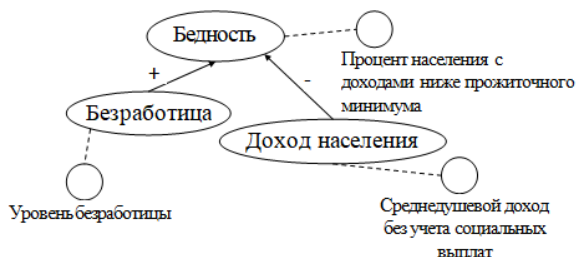


Рис.1 – Факторы тестовой когнитивной модели

Система нечеткого вывода получена алгоритмом равномерного разбиения входного пространства [4]. По двум факторам для всех объектов были получены оценки по треугольным функциям принадлежности – «L» для низкого уровня фактора и «H» для высокого уровня. На основе каждой из 4 комбинаций нечетких оценок были получены продукционные правила. Консеквенты правил были получены путем построения регрессионных моделей на основе выборки из объектов, степень принадлежности которых к группе равна или превышает 0.5. Если количество элементов в выборке менее 20, в выборку добавлялись элементы с наиболее близкими к данной группе значениями.

Результатом стали следующие правила:

$$R_1: \text{ЕСЛИ } F_2 = \langle L \rangle \text{ И } F_3 = \langle L \rangle, \text{ ТО } y(x_1, x_2) = 0.47x_1 - 0.84x_2 + 0.56$$

$$R_2: \text{ЕСЛИ } F_2 = \langle L \rangle \text{ И } F_3 = \langle H \rangle, \text{ ТО } y(x_1, x_2) = -0.37x_1 + 0.03x_2 + 0.36$$

$$R_3: \text{ЕСЛИ } F_2 = \langle H \rangle \text{ И } F_3 = \langle L \rangle, \text{ ТО } y(x_1, x_2) = -0.04x_1 - 0.99x_2 + 0.91$$

$$R_4: \text{ЕСЛИ } F_2 = \langle H \rangle \text{ И } F_3 = \langle H \rangle, \text{ ТО } y(x_1, x_2) = 0.12x_1 - 0.47x_2 + 0.55$$

Далее были спрогнозированы значения для оценки фактора  $F_1$ . Характеристики качества предсказания полученных моделей можно увидеть в таблице 2. Были использованы коэффициент детерминации ( $R^2$ ), средние абсолютная (MAE) и квадратичная (MSE) ошибки.

Таблица 2 – Качество результатов прогнозирования

Метод	$R^2$	MAE	MSE
Линейная регрессия	0.67	2.22	7.34
Нечеткий вывод	0.84	1.50	3.51

Можно увидеть, что использование нечеткого вывода в гибридной когнитивной модели позволяет улучшить точность получаемых прогнозов благодаря учету принадлежности объектов к различным группам. Далее стоит рассмотреть возможности статического анализа когнитивной карты с помощью полученной базы правил. Равномерным разбиением можно получить нечеткие оценки для выходного фактора  $F_1$ . У каждого региона в 2021 году для прогноза и реальной числовой оценки были получены принадлежности термам: «L» для низкого, «M» для среднего и «H» для высокого уровня. Также с помощью нечеткого вывода можно получить специфичные для региона значения силы влияния. Результаты представлены в таблице 3. Для различных регионов обнаруживаются существенно отличающиеся сила и направление влияния факторов, для некоторых регионов эти оценки значительно отличаются от оценок, полученных простой регрессионной моделью.

**Таблица 3 – Оценки элементов карты для региона**

Регион	Оценка $F_1$ (прогноз)	Оценка $F_1$ (реальная)	Влияние $F_2$ на $F_1$	Влияние $F_3$ на $F_1$
Алтайский край	M	M	0.26	-0.68
Иркутская область	M	M	0.12	-0.55
Кемеровская область	M	L	0.20	-0.61
Красноярский край	M	M	-0.15	-0.19
Новосибирская область	L	L	-0.08	-0.34
Омская область	M	L	0.08	-0.54
Республика Алтай	H	H	0.05	-0.90
Республика Тыва	H	H	-0.04	-0.99
Республика Хакасия	M	M	0.23	-0.69
Томская область	M	L	0.07	-0.54

Использование нечеткого вывода не только помогло улучшить возможности прогнозирования с гибридной когнитивной картой, но и позволило получить детальные оценки для каждого отдельного объекта с учетом его конкретной специфики. Такие оценки могут быть полезны и для анализа статического состояния конкретного объекта, и для наблюдения его динамики во времени. Можно делать выводы о поведении отдельных групп объектов, об их уникальных свойствах. В будущем планируется рассмотреть более сложные алгоритмы группировки данных и выделения силы влияния для улучшения качества баз правил и точности прогнозирования.

Литература:

1. Силич М.П., Силич В.А., Аксенов С.В. Способы оценки факторов иерархической гибридной когнитивной карты на примере оценки состояния теплопотребления регионов // Проблемы управления. – 2018. – № 1. – С. 58–65.
2. Carvalho J.P., Tome J. Rule based fuzzy cognitive maps and fuzzy cognitive maps—A comparative study // 18th International Conference of the North American Fuzzy Information Processing Society – New York, 1999. – P. 115–119.
3. Росстат – Официальная статистика [Электронный ресурс] // Федеральная служба государственной статистики – URL: <https://rosstat.gov.ru/folder/10705> (дата обращения: 01.11.2023)
4. Сарин К.С., Ходашинский И.А. Три алгоритма генерации структуры нечеткой системы типа Такаги-Сугено // Знания - Онтологии - Теории (ЗОНТ-2015): Материалы Всероссийской конференции с международным участием, Новосибирск, 06–08 октября 2015 года – Российская академия наук, Сибирское отделение; Институт математики им. С.Л. Соболева. Том 2. – Новосибирск: ООО "Технотрейд", 2015 – С. 124–132.

## **МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РАСПРЕДЕЛИТЕЛЬНОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ**

**А.С. Росущан**

**Национальный исследовательский университет  
«Высшая школа экономики»,  
г. Москва, [alexander.rosushchan@outlook.com](mailto:alexander.rosushchan@outlook.com)**

*В статье выполнена работа по моделированию сверточной нейронной сети на языке программирования Python. В процессе обучения алгоритмом машинного обучения (ML) она распознала и предсказала с определенной вероятностью представленные геометрические фигуры.*

*The article performed work on modeling a convolutional neural network in the Python programming language. It has successfully recognized and predicted the presented geometric shapes with a certain probability, after the process of training with a machine learning (ML) algorithm.*

Сегодня изучение искусственного интеллекта (ИИ) является крайне актуальным направлением, так как позволяет упрощать или более

эффективно реализовывать различные процессы, что может быть полезно для ряда сфер от бизнеса, до медицины и промышленности. Нейронные сети позволяют создавать самообучающуюся систему, которая получает входные данные, обрабатывает и анализирует их, а по итогу даёт выходные данные, при этом обучаясь на каждой итерации. Конкретно технология распознавания образов используется в маркетинге (поиск товаров определенного бренда), медицине (определение заболеваний по рентгеновскому снимку), машиностроение (фотосепараторы, системы автопилотирования) и тд. Таким образом, тему необходимо изучать, так как при помощи методов распознавания изображений можно уменьшить роль человека в случаях, где важно быстрое реагирование на происшествия. Разработка алгоритма распознавания изображений исправит несовершенства существующих систем, сократит использование ресурсов и затраченное время.

Предмет исследования – математическое моделирование распределительной системы распознавания образом с помощью сверточной нейронной сети. Целью работы является создание модели для классификации изображений геометрических фигур с помощью сверточных нейронных сетей (СНС) [1]. Для достижения данной цели используются язык программирования Python (библиотеки Tensorflow и Keras), а также Microsoft Visio для создания блок-схем (Рис. 1) и Microsoft PowerPoint для создания визуализаций (Рис .2).

При разработке программного обеспечения наряду с решением математических практических задач, все чаще используются алгоритмы машинного обучения обработки изображений с использованием нейронных сетей для их распознавания. Сверточная нейронная сеть использует математические операции свертки, благодаря чему удается упорядочить процедуру получения цифровой информации из них.

Абстрактно работу виртуальной машины можно представить как математическую модель работы мозга человека. В случае виртуальной машины мы имеем перцептрон, а в случае человека биологическую сеть из нейронов[2]. В полносвязных сетях прямого распространения (FFNN) нейроны предыдущего слоя связаны с нейронами следующего слоя. Распространение сигнала происходит от первого (входного) к последнему слою (выход сети) без обратных связей между элементами сети (см. Рис. 1).

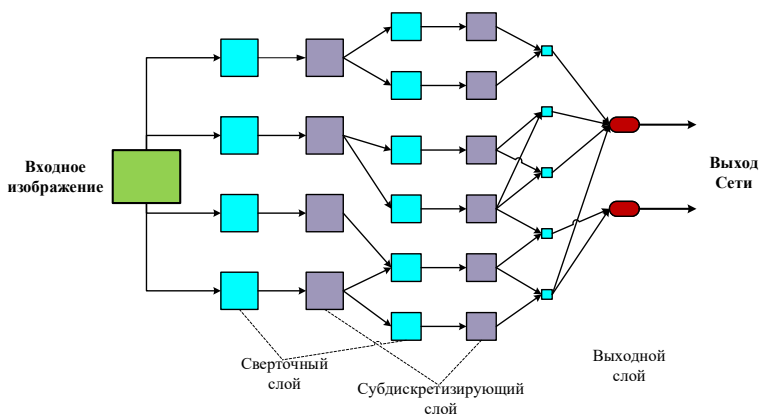


Рис. 1 – Архитектура сверточной нейронной сети

Связи между нейронами имеют определенные веса от  $i$ -го нейрона к  $j$ -му нейрону. Каждый нейрон получает сумму входных значений, умноженных на вес соответствующей связи

$$Sum = \sum_{k=1}^N w_{ij} \cdot x_i, \quad (1)$$

где  $w_{ij}$  – вес связи от  $i$ -го нейрона к  $j$ -му нейрону;  $x_i$  – значение входного сигнала [6]. Сумма, полученная после вычисления, передается функции активации – специальной математической формуле, которая активирует или не активирует очередной нейрон.

Можно выделить несколько типов нейронных сетей:

1. С прямой связью (feedforward neural network, FF или FFNN).
2. Рекуррентная нейронная сеть (recurrent neural network, RNN).
3. Сверточная нейронная сеть (convolutional neural network, CNN).

Преимущества CNN связаны, во-первых, с быстрым обучением, во-вторых, с высоким уровнем эффективности во время практического применения при работе по поиску шаблонов, распознаванию образов, и даже работе с аудиофайлам [6].

Принцип вычисления сверточной нейронной сети может быть представлен как матрица входных данных (5x5), которая сканируется матрицей ядра (3x3). Благодаря поэлементному умножению части исходной матрицы и суммированием всех полученных значений в один выходной пиксель и получается ядро [8].

Следующим этапом алгоритма является фильтрация. Каждый фильтр обрабатывает вход со своим отличающимся от других набором ядер и скалярным смещением и, по итогу, создает один выходной канал [3].

После процесса фильтрации происходит обучение сети. Параметры СНС настраиваются посредством математического моделирования среды, в которую встроена нейросеть. Способ подстройки параметров определяет тип обучения, существует два типа обучения (с учителем и без учителя). Для данной работы выбрано обучение с учителем, т. е. обучение с упорядоченной выборкой обучающих примеров (см. Рис 2).

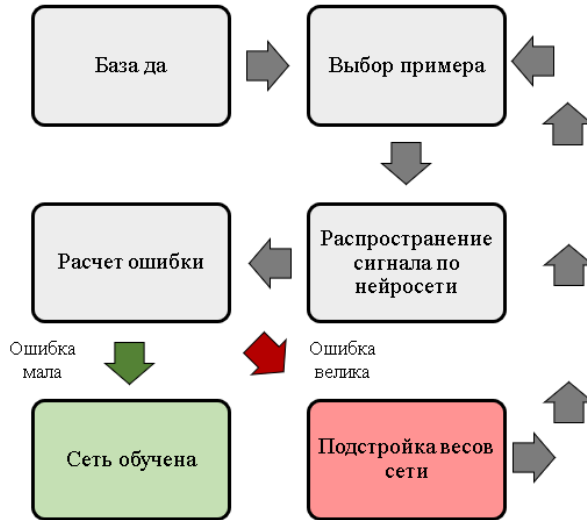


Рис. 2 – Схема обучения Сверточной нейронной сети

Для этого разные обучающие образцы подаются на входы сети, обрабатываются, а после выходной сигнал сравнивается с образцовым значением целевого вектора [4]. Ошибка отклонения от образца приводит к изменению весовых коэффициентов связей внутри сети в зависимости от выбранного алгоритма машинного обучения.

Для осуществления математического моделирования с помощью CNN – распознавания геометрических фигур используем язык программирования Python. Однородные традиционные данные для работы с CNN записываются в датасет (*dataset*), как показано на рис. 3. Например, большие базы данных (*Big Data*) имеют разнородную (числа, текст, изображения, аудио, видео и т.п.) информацию.

Разработка нейронной сети выполняется в облачном сервисе *Google Colaboratory* на языке *Python* с добавлением библиотек *Tensorflow* и *Keras*. Далее создадим переменные, которые будут содержать названия наших классов:

```

target_name = ["triangle", "rhombus", "square"] dir_dataset =
"dataset_figures"
train_dir = PATH + dir_dataset + '/train' test_dir = PATH + dir_dataset
+ '/test' pred_dir = PATH + dir_dataset + '/pred'
dir_dataset - корневой каталог для всех имеющихся на
диске датасетов. Переменные train_dir, test_dir, pred_dir - это пути,
которые указывают на каталоги с данными для: обучения, тестирования,
предсказания.
CNN работает с данными постепенно, обработка происходит
последовательно (частями).
img_width, img_height = 32, 32 input_shape = (img_width, img_height,
1) batch_size = 128
N = len(target_name)
print ("количество видов в классификации " + str(N))

```

Прокомпилировав часть скрипта, получается выход данных о том, что в данной сети 4 вида классификации (ромбы, треугольники, квадраты, окружности).

Обучение нейронной сети проводим от однослойного до трехслойного сетей, для того чтобы уменьшился процент ошибок распознавания.

Меняем размерность, добавляем фильтры. Количество эпох = 25. Далее создаем сверточную нейронную сеть (СНС).

С помощью функции summary просматриваются все характеристики полученной СНС модели. Можно заметить, что у модели всего 100 418 параметров и все они являются тренируемыми.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 10, 10, 32)	320
max_pooling2d (MaxPooling2D)	(None, 5, 5, 32)	0
conv2d_1 (Conv2D)	(None, 3, 3, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 1, 1, 64)	0
flatten (Flatten)	(None, 64)	0
dense_5 (Dense)	(None, 1200)	78000
dense_6 (Dense)	(None, 3)	3603
-----		
Total params: 100,418		
Trainable params: 100,418		
Non-trainable params: 0		

Рис. 3 – Характеристики Сверточной нейронной сети



По итогу обучения мы имеем следующие метрики модели: точность – 1, точность в проверочном наборе – 0,8087 (см. Рис 4).

```

11/11 [-----] - 0s 22ms/step - loss: 0.0182 - accuracy: 0.9978 - val_loss: 1.4427 - val_accuracy: 0.7716
Epoch 14/27
11/11 [-----] - 0s 21ms/step - loss: 0.0169 - accuracy: 0.9971 - val_loss: 1.3682 - val_accuracy: 0.7971
Epoch 15/27
11/11 [-----] - 0s 22ms/step - loss: 0.0116 - accuracy: 0.9993 - val_loss: 1.4661 - val_accuracy: 0.7739
Epoch 16/27
11/11 [-----] - 0s 21ms/step - loss: 0.0091 - accuracy: 0.9993 - val_loss: 1.4589 - val_accuracy: 0.7913
Epoch 17/27
11/11 [-----] - 0s 24ms/step - loss: 0.0075 - accuracy: 0.9993 - val_loss: 1.4520 - val_accuracy: 0.8058
Epoch 18/27
11/11 [-----] - 0s 25ms/step - loss: 0.0069 - accuracy: 1.0000 - val_loss: 1.4897 - val_accuracy: 0.7942
Epoch 19/27
11/11 [-----] - 0s 20ms/step - loss: 0.0083 - accuracy: 0.9985 - val_loss: 1.5117 - val_accuracy: 0.8887
Epoch 20/27
11/11 [-----] - 0s 20ms/step - loss: 0.0056 - accuracy: 1.0000 - val_loss: 1.5047 - val_accuracy: 0.7971
Epoch 21/27
11/11 [-----] - 0s 24ms/step - loss: 0.0030 - accuracy: 1.0000 - val_loss: 1.5208 - val_accuracy: 0.8858
Epoch 22/27
11/11 [-----] - 0s 27ms/step - loss: 0.0031 - accuracy: 1.0000 - val_loss: 1.5054 - val_accuracy: 0.8058
Epoch 23/27
11/11 [-----] - 0s 25ms/step - loss: 0.0027 - accuracy: 1.0000 - val_loss: 1.5301 - val_accuracy: 0.8058
Epoch 24/27
11/11 [-----] - 0s 25ms/step - loss: 0.0025 - accuracy: 1.0000 - val_loss: 1.5500 - val_accuracy: 0.8058
Epoch 25/27
11/11 [-----] - 0s 23ms/step - loss: 0.0024 - accuracy: 1.0000 - val_loss: 1.5458 - val_accuracy: 0.8145
Epoch 26/27
11/11 [-----] - 0s 25ms/step - loss: 0.0024 - accuracy: 1.0000 - val_loss: 1.5677 - val_accuracy: 0.8087
Epoch 27/27
11/11 [-----] - 0s 26ms/step - loss: 0.0020 - accuracy: 1.0000 - val_loss: 1.5709 - val_accuracy: 0.8887
    
```

Рис. 4 – Ход обучения Сверточной нейронной сети

Нейронная сеть достигает максимума в позиции 0,8707 (см. Рис. 5). Это говорит о том, что нейронная сеть прекращает обучение, результативность процесса обучения составляет приблизительно 87 %.

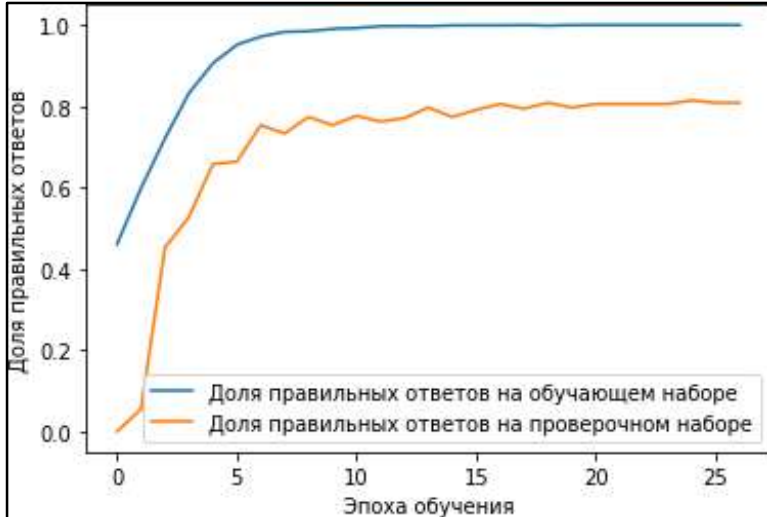


Рис. 5 – Динамика правильных ответов во время обучения нейронной сети

На рисунке 6 представлен распознанный и истинный объект «треугольник» в ходе работы СНС. Также нейронная сеть делает предсказание в виде подсчета вероятности.

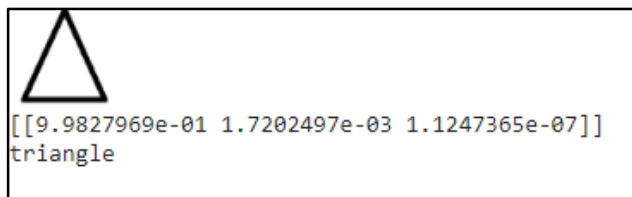


Рис. 6 – Характеристики объекта «треугольник»

На рисунке 7 показано отображение подсчета вероятности. Первое числовое значение ( $9,9981445e-01$ ) говорит о вероятности отношения данного объекта к треугольникам, второе ( $1,8349843e-04$ ) к ромбам, третье значение ( $2,0511375e-06$ ) – к квадратам. Выбирая наибольшее значение вероятности, определяется классификация фигуры.

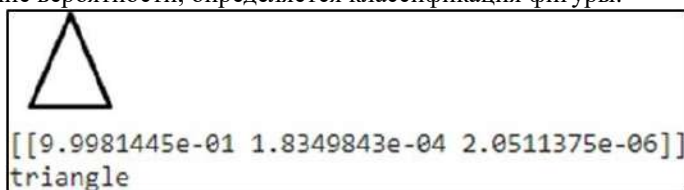


Рис. 7– Объект предсказания нейронной сети

Подводя итог, в данном исследовании был представлен обзор видов нейронных сетей, а также описан принцип работы сверточной нейронной сети. В ходе выполнения работы была смоделирована сверточная нейронная сеть на языке Python. В процессе обучения алгоритмом машинного обучения она распознала и предсказала с определенной вероятностью представленные геометрические фигуры (треугольник, квадрат, ромб). Результативность нейронной сети составила 87%.

#### Литература:

1. *Буйначев С.К., Боклаг Н.Ю.* Основы программирования на языке Python : учебное пособие . – Екатеринбург: Изд-во Уральского университета, 2021. – 92 с.
2. *Гласснер Э.* Глубокое обучение без математики. Том 2. Практика: руководство / перевод с англ. В. А. Яроцкого. – М.: ДМК Пресс, 2020. – 610 с.
3. *Душкин Р. В.* Искусственный интеллект. – М.: ДМК Пресс, 2019. – 280 с.
4. *Ибрагим А. С.* Сравнительный анализ искусственных нейронных сетей и мультиагентных систем // Известия Кабардино-Балкарского научного центра РАН. – 2021. – №1 (99). – С. 20-27.

5. Кабанова В. В. Генерация и трансформация лица на изображении с использованием искусственного интеллекта / В. В. Кабанова // Регион искусственного интеллекта : Материалы Всероссийской студенческой научной конференции, Череповец, 06 ноября 2022 года. – Череповец: Череповецкий государственный университет, 2022. – С. 33-37.

6. Кабанова В. В. Применение искусственного интеллекта при работе с мультимедийной информацией / В. В. Кабанова, О. С. Логунова // Вестник Череповецкого государственного университета. – 2022. – № 6(111). – С. 23-41.

7. Романюк А. Г., Смирнов А. Н., Антонова В. М. Использование глубокого обучения нейросети для распознавания голосовых команд пользователя. // Журнал радиоэлектроники. – 2019. – №11. С. 11.

8. Исследование сверточных нейронных сетей для обнаружения объектов на аэрокосмических снимках / Скрипачев В. О., Гуйда М. В., Гуйда Н. В., Жуков А. О. // International Journal of Open Information Technologies. – 2022. – №7. – С. 54-60

## **О МЕТОДАХ КЛАСТЕРИЗАЦИИ ДИНАМИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ**

**Д.В. Рыжков**

**Новосибирский государственный университет,  
г. Новосибирск, d.ryzhkov@g.nsu.ru**

**Научный руководитель: Ломов А.А., д.ф.-м.н., доцент**

*В работе рассматривается задача кластеризации одномерных временных рядов. Используемый подход к данной задаче заключается в том, что временным рядам сопоставляются описывающие их динамические системы и кластеризация проводится уже в пространстве динамических систем. В работе предложены модификации класса спектральных метрик динамических систем, необходимых для проведения кластеризации.*

*The paper considers the problem of clustering of one-dimensional time series. The approach to this problem is that time series are matched with dynamical systems describing them and clustering is carried out in the space of dynamical systems. In this paper we propose modifications of spectral metrics of dynamical systems necessary for clustering.*

Кластеризация представляет собой задачу разбиения множества наблюдений на не пересекающиеся подмножества таким образом, чтобы элементы одного подмножества были близки друг к другу, а элементы разных подмножеств были далеки друг от друга. Рассматриваемые нами

множества наблюдений являются множествами временных рядов. Данные временных рядов представляют собой последовательности измерений во времени, описывающие поведение некоторой систем.

Из трех основных подходов к кластеризации временных рядов [1] нами был выбран подход, который заключается в том, что мы сопоставляем временным рядам динамические системы типа АРСС

$$y_{k+n} + \alpha_{n-1}y_{k+n-1} + \dots + \alpha_0 y_k = \beta_m u_{k+m} + \dots + \beta_0 u_k, \quad k \in \mathbb{Z}, \quad (1)$$

где  $\{y(k)\}_{k \in \mathbb{Z}} \in \mathbb{R}$  – временной ряд,  $\{u(k)\}_{k \in \mathbb{Z}} \in \mathbb{R}$  – входной сигнал динамической системы (считаем, что он представляет собой белый гауссовский шум с дисперсией равной  $\sigma^2$ , т.е.  $u(k) \sim \mathcal{N}(0, \sigma^2)$ ,  $k \in \mathbb{Z}$ ).

Затем кластеризация проводится уже для динамических систем. То есть задача кластеризации временных рядов при таком подходе сводится к задаче кластеризации динамических систем.

В данной работе мы будем рассматривать следующую постановку задачи кластеризации – задачу  $k$ -медоидов.

$$\begin{cases} \sum_{i=1}^n \sum_{x_j \in S_i} d(c_i, x_j) \rightarrow \min_{S_1, \dots, S_l} \\ S_1 \cup \dots \cup S_l = \{x_1, \dots, x_N\}, \quad S_i \cap S_j = \emptyset, \quad i \neq j, \end{cases}$$

где  $x_1, \dots, x_N$  – выборка,  $S_1, \dots, S_l$  – кластеры,  $c_1, \dots, c_l$  – медоиды соответствующих кластеров. Предполагается, что количество кластеров  $l$  известно.

Медоид  $c_i$  кластера  $S_i$  – элемент выборки, принадлежащий кластеру  $S_i$ , такой что сумма расстояний от него до остальных элементов кластера  $S_i$  минимальна, т.е.  $c_i = \arg \min_{x_j \in S_i} \sum_{x_k \in S_i} d(x_j, x_k)$ .

Для решения задачи кластеризации динамических систем необходимо определить метрику  $d(\cdot, \cdot)$  в пространстве динамических систем. Рассматриваемые нами метрики определяются с помощью передаточной функции  $H(z)$  динамической системы. Передаточная функция  $H(z)$  системы (1) представляет собой несократимую рациональную функцию

$$H(z) = \frac{\beta_m z^m + \dots + \beta_0}{z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_0} = \frac{\beta(z)}{\alpha(z)}.$$

Корни  $z_1, \dots, z_m$  числителя  $\beta(z)$  и корни  $p_1, \dots, p_n$  знаменателя  $\alpha(z)$  называем, соответственно, нулями и полюсами передаточной функции.

Введем, вслед за авторами [2], в рассмотрение класс спектральных метрик динамических систем, определяемых следующим образом:

$H_2$ -расстояние ( $\mathbf{h}_2$ )

$$d_{h_2}(H^{(1)}, H^{(2)}) = \|H^{(1)} - H^{(2)}\|_{h_2}, \quad \|H(z)\|_{h_2}^2 = \frac{1}{2\pi} \int_0^{2\pi} |H(e^{i\omega})|^2 d\omega,$$

$H_\infty$ -расстояние ( $\mathbf{h}_\infty$ )

$$d_{h_\infty}(H^{(1)}, H^{(2)}) = \|H^{(1)} - H^{(2)}\|_{h_\infty}, \quad \|H(z)\|_{h_\infty} = \sup_{\omega \in [0, 2\pi]} |H(e^{i\omega})|,$$

Расстояние между спектрами ( $\mathbf{sp}$ )

$$d_{sp}^2(H^{(1)}, H^{(2)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\Phi^{(1)}(\omega) - \Phi^{(2)}(\omega))^2 d\omega,$$

где  $\Phi^{(1)}$  и  $\Phi^{(2)}$  – спектральные плотности мощности сигналов.

В работе предложены модификации спектральных метрик, заключающиеся в применении параметрического преобразования  $f^*(H; q)$  к передаточным функциям динамических систем, перед вычислением расстояния между ними. Предложенные модифицированные метрики имеют следующий вид:

$H_2^*$ -расстояние ( $\mathbf{h}_2^*$ )

$$d_{h_2^*}(H^{(1)}, H^{(2)}) = d_{h_2}(f^*(H^{(1)}), f^*(H^{(2)})), \quad f^*(H; q) = \frac{H}{q + H},$$

$H_\infty^*$ -расстояние ( $\mathbf{h}_\infty^*$ )

$$d_{h_\infty^*}(H^{(1)}, H^{(2)}) = d_{h_\infty}(f^*(H^{(1)}), f^*(H^{(2)})),$$

Расстояние между спектрами ( $\mathbf{sp}^*$ )

$$d_{sp^*}(H^{(1)}, H^{(2)}) = d_{sp}(f^*(H^{(1)}), f^*(H^{(2)})).$$

Нами были исследованы некоторые аналитические свойства предложенного преобразования. В частности, были доказаны утверждения, формулирующие достаточные условия на выбор параметра  $q$ , при которых предложенное преобразование  $f^*(H; q)$  сохраняет свойства устойчивости и минимально-фазовости передаточной функции. Данные свойства используются при численном подсчете спектральных метрик [3-4].

В работе был проведен ряд численных экспериментов, в которых сравнивались предложенные модификации спектральных метрик с их изначальными версиями. Для определения качества полученной кластеризации использовался скорректированный индекс Ранда – это

внешняя мера качества кластеризации, основанная на сравнении результата кластеризации с априори известным разделением на классы. Скорректированный индекс Ранда имеет значение, близкое к 0 (возможно отрицательное), когда полученное разбиение на кластеры сопоставимо со случайным и означает, что полученное разбиение на кластеры недостоверно, и равен 1, когда кластеры совпадают с истинными.

В качестве примера в таблице 1 приведены результаты одного из проведенных экспериментов, в котором проводилась кластеризация временных рядов, соответствующих динамическим системам 1-го порядка с одинаковыми фиксированными полюсами и различными нулями. В ячейках таблицы представлены средние значения скорректированных индексов Ранда, полученные при 10-кратном повторении эксперимента.

**Таблица 1 – Результаты численного эксперимента**

Расстояние между нулями	$h_2$	$h_2^*$	$h_\infty$	$h_\infty^*$	sp	sp*
0.1	0.024	0.221	0.021	0.187	0.027	0.151
0.15	0.049	0.408	0.037	0.418	0.043	0.385
0.2	0.141	0.561	0.080	0.548	0.088	0.501
0.25	0.292	0.882	0.101	0.801	0.142	0.717

Как мы можем увидеть из результатов, модифицированные спектральные метрики  $h_2^*$ ,  $h_\infty^*$  и  $sp^*$  показали себя лучше своих изначальных версий  $h_2$ ,  $h_\infty$  и  $sp$ . Спектральные метрики  $h_2$ ,  $h_\infty$  и  $sp$ , показали результат, сопоставимый со случайным распределением объектов по кластерам.

В итоге, в работе предложены модификации метрик динамических систем для решения задачи кластеризации временных рядов. Сформулированы и доказаны утверждения об аналитических свойствах предложенных метрик и проведены численные эксперименты, в которых показано, что модификация не ухудшает, а в некоторых случаях значительно улучшает качество кластеризации. Экспериментально исследовано влияние параметра модифицированных метрик на результат кластеризации и даны рекомендации по выбору параметра для конкретных наборов данных.

Литература:

1. Rani S., Sikka G. Recent Techniques of Clustering of Time Series Data: A Survey // International Journal of Computer Applications. – 2012. – Vol. 52, N 15. – P. 1–9.

2. Boets J., De Cock K., De Moor B. Distances Between Dynamical Models For Clustering Time Series // IFAC Proceedings Volumes. – 2006. – Vol. 39, N 1. – P. 392–397.

3. Bruinsma N., Steinbuch M. A fast algorithm to compute the  $H_\infty$ -norm of a transfer function matrix // Systems & Control Letters. – 1990. – Vol. 14, N 4. – P. 287–293.

4. Zhou K., Doyle J. C., Glover K. Robust and Optimal Control. – USA: Prentice-Hall, Inc., 1996. – 596 p.

## **ОБРАБОТКА ДАННЫХ ПОДВОДНОГО АКУСТИЧЕСКОГО ВОЛНОГРАФА В ХОДЕ ДОЛГОСРОЧНОГО НАТУРНОГО ЭКСПЕРИМЕНТА В ЧЕРНОМ МОРЕ**

**М.С. Рябкова, К.С. Епанова, Ю.А. Титченко**

**Институт прикладной физики РАН,**

**г. Нижний Новгород, [mrjabkova@ipfran.ru](mailto:mrjabkova@ipfran.ru)**

**Научный руководитель: Караев В.Ю., к.ф.-м.н., зав. лаб.**

*Подводный акустический волнограф предоставляет уникальные возможности для всепогодного мониторинга водной поверхности. Антенна локатора ориентирована на водную поверхность, локатор использует ультразвуковые импульсы для сканирования поверхности. В работе представлены два метода измерения высоты значительного волнения, показано, что они имеют высокий коэффициент корреляции.*

*The underwater acoustic wave gauge provides unique capabilities for all-weather monitoring of the water surface. The locator antenna is oriented towards the water surface, and the locator uses ultrasonic pulses to scan the surface. Two methods for measuring significant wave height are presented in the study, demonstrating a high correlation coefficient between them.*

В настоящее время основную информацию о верхнем слое Мирового океана получают благодаря космической радиолокации. Однако эффективное использование радиолокационных данных невозможно без постоянной проверки и подтверждения точности восстанавливаемой "тематической" информации. Для этого активно применяются контактные измерения, получаемые, например, с использованием волнографов и морских буев, размещенных в различных районах Мирового океана [1]. Данные от буев широко используются для проверки и подтверждения правильности алгоритмов обработки

радиолокационных данных. Однако применение буев ограничено в замерзающих водоемах и в периоды сильных штормов.

Для постоянного мониторинга ветрового волнения и ледяного покрова предлагается использовать подводный акустический волнограф. Прибор устанавливается вблизи дна, он оснащен импульсным и доплеровским гидролокаторами, работающими на частоте 200 кГц. Антенна импульсного гидролокатора ориентирована вертикально вверх, на водную поверхность, в то время как антенна доплеровского гидролокатора отклонена на 5 градусов от направления в зенит. Импульсный гидролокатор излучает акустические импульсы длиной от 5 до 40 мкс с частотой повторений от 15 до 100 Гц. Вариант прибора, предназначенный для использования в южных морях, может работать на глубинах до 100 метров, в то время как версия для арктических морей способна выдерживать погружение на глубину до 300 метров.

Акустический волнограф производил непрерывные измерения на полигоне «Геленджик» ЮО ИО РАН на Черном море в течение 10 месяцев с декабря 2019 по сентябрь 2020 года. Прибор был установлен на подводной платформе, расположенной на глубине 28 м в 1 км от берега, он управлялся дистанционно и каждый час в течение 15 минут производил запись.

Измеряя время прихода максимального отражения, можно определить расстояние от волнографа до поверхности воды, и на основе зависимости расстояния от прибора до водной поверхности по времени можно вычислить энергетический спектр волнения. Анализ этого спектра позволяет определить различные характеристики волнения. В работе [3] этот метод описан более подробно и показано, что точность измерения высоты волнения с помощью акустического волнографа сопоставима с точностью ADCP.

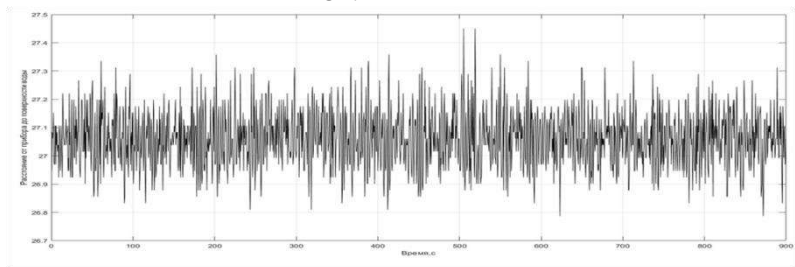


Рис.1 – Зависимость расстояния от акустического волнографа до поверхности воды от времени (06-08-2020 01:48:34)



Применяя преобразование Фурье к зависимости расстояния от волнографа до поверхности воды от времени можно получить спектр волнения.

Зная спектр волнения  $S(f)$ , характеризующий распределение распределение энергии по волнам с различной частотой  $f$ , можно вычислить, например, высоту значительного волнения (средняя высота волны самой высокой трети волн):

$$H_s = 4\sqrt{\int S(f)df} \quad (1)$$

Можно усреднить все отраженные импульсы, полученные за 15 минут работы прибора и проанализировать полученный средний импульс. В работе [3] показано, что передний склон такого импульса определяется значительной высотой волнения и глубиной погружения прибора.

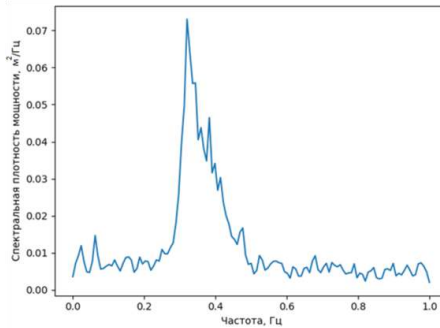


Рис.2 – Спектр волнения, полученный для записи на рис. 1.

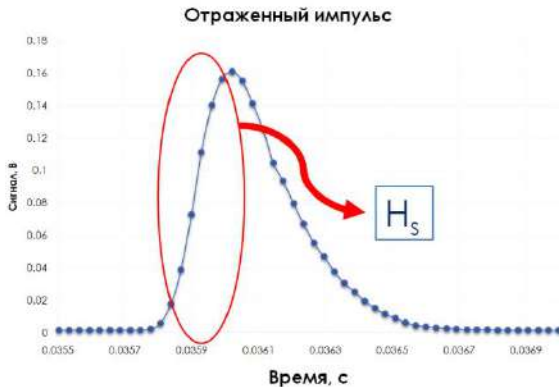


Рис.3 – Усредненный отраженный импульс, полученный для записи на рис. 1.

Было проведено сравнение этих методов для всех измерений, полученных с 2.12.2019 по 15.09.2020. На рисунке 4 представлены измерения высоты значительного волнения, полученного разными методами.

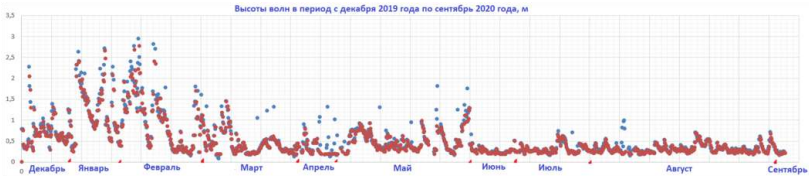


Рис.4 – Высота значительного значения (в м), рассчитанная методом вычисления спектра (синие точки) и методом анализа усредненного импульса (красные точки).

Коэффициент корреляции двух методов составляет 0,97. Также оба метода позволяют установить уровень воды (среднее расстояние от прибора до поверхности воды за 15 минут записи), для этого параметра коэффициент корреляции составляет 0,9.

Акустический волнограф предоставляет уникальные возможности для всепогодного мониторинга водной поверхности, он может работать долгое время и непрерывно поставлять данные о поверхностном волнении. Два метода измерения высоты значительного волнения с помощью акустического волнографа имеют высокий коэффициент корреляции.

#### Литература:

1. Коровин В.П., Океанологические наблюдения в прибрежной зоне моря. Учебное пособие. – СПб.: РГГМУ, 2007. – 434 с.
2. Рябкова М.С., Титченко Ю.А., Караев В.Ю., Мешков Е.М., Беляев Р.В., Яблоков А.А., Баранов В.И., Очередник В.В. Измерение статистических характеристик морской поверхности с помощью подводного акустического волнографа в Чёрном море и сравнение с ADCP // Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. №2. С. 189–204 . – 2021. – Т. 18. – С. 189–204.
3. Караев В. Ю., Мешков Е. М., Титченко Ю. А. Подводный акустический высотомер // Известия ВУЗов, сер. Радиофизика. – 2014. – Т. 57. – С. 543–554.

# ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРЕДСКАЗАНИЯ ОТТОКА АБОНЕНТОВ ИНТЕРНЕТ-ПРОВАЙДЕРА

А.С. Сухих

Новосибирский государственный технический университет,  
г. Новосибирск, [suxix.2018@stud.nstu.ru](mailto:suxix.2018@stud.nstu.ru)

Научный руководитель: Стасьшин В.М., к.т.н., доцент

*Предсказание оттока абонентов российского интернет-провайдера как решение задачи бинарной классификации с применением моделей машинного обучения. Используются модели наивного Байесовского классификатора, дерева решений, логистической регрессии, метода опорных векторов и случайного леса. В качестве наилучшей модели было выбрано дерево решений. Результаты обученной модели на тестовых данных имеют Precision, Recall и F1-score равные 0.84.*

*Predicting customer churn for a Russian Internet service provider (ISP) is a binary classification problem that can be solved using machine learning. Naive Bayes classifier, decision tree, logistic regression, support vector machine and random forest models were trained and tested on an ISP dataset. Decision Tree showed the best Precision, Recall and F1-score results - 0.84.*

Вопрос оттока в сфере телекоммуникаций обстоит достаточно остро, поскольку в данной области крайне высокий уровень конкуренции. Также характерно то, что стоимость удержания уже имеющихся абонентов существенно выше, чем стоимость привлечения и подключения новых. Чаще всего компании начинают удержание клиента тогда, когда он уже сам сообщил о своем намерении прекратить пользование услугами. Возможность предсказать, какие из клиентов скорее всего захотят расторгнуть договор с компанией, позволит снизить отток, предприняв превентивные меры по удержанию клиентов.

Применение методов машинного обучения позволит установить явные и неявные зависимости, влияющие на отток абонентов, и с некоторой точностью предсказать, какие клиенты вероятнее всего прекратят использование услуг.

Поскольку основной целью является отнесение абонентов к одному из классов: лояльных к компании или склонных к уходу, то система должна решать задачу бинарной классификации. Широкое применение в решении задач классификации нашли различные методы машинного обучения. В ряде работ [1,5,6] был использован наивный Байесовский классификатор (Naive Bayes). Также частым выбором является дерево решений (Decision Tree) [1,3,4,5,6] и его ансамблевая модификация «случайный лес» (Random Forest) [2,5,6]. Кроме того, в некоторых

работах используется метод опорных векторов (Support Vector Machine) [1,2,3] и другие. Приведенные классификаторы были использованы в данной работе для обучения модели, наиболее точно предсказывающей отток абонентов. Для сравнения результатов работы моделей были использованы метрики Accuracy, Precision, Recall и F1-score.

Для составления датасета абонентов, на основе которого выполнялось обучение моделей, интернет-провайдером было предоставлено хранилище данных, аккумулирующее в себе данные различных эксплуатационных систем, таких как CRM-система, биллинговая система, телефония технической поддержки и прочие.

В большинстве случаев интернет-провайдеры предоставляют услуги в виде подписок с некоторой фиксированной абонентской платой. Для такого рода услуг закономерным событием возникновения оттока стоит считать прекращение подписки. Также были выделены следующие группы признаков:

- общая информация об абоненте и услугах (пол, возраст, срок жизни услуги, наличие других услуг);
- информация о подключении (район подключения, модель коммутатора, количество других операторов в доме);
- обращаемость (количество обращений за период, количество заявок по абоненту);
- платежи и блокировки (количество и размер платежей за различные периоды, блокировки лицевого счета и т. д.).

С помощью SQL-запросов к хранилищу данных был составлен датасет с 48 признаками за некоторый период времени. Датасет содержит в себе два типа абонентов: прекративших использование услуг в заданном периоде и имеющих активную услугу. Поскольку состояние активных абонентов в заданном периоде может быть определено сразу в множестве моментов времени, вычисление признаков выполнялось в случайный момент времени для каждого абонента. Это позволило избежать смещения значений некоторых параметров, связанных с расчетом на одном конкретном моменте времени рассматриваемого периода.

Изначальный датасет имеет сильный дисбаланс классов в пользу класса «нет оттока», который был устранен процедурой undersampling, заключающейся в исключении случайных наблюдений большего класса до равенства объемов между классами. После undersampling количество строк в датасете составило 45 870.

Модель из 48 параметров достаточно сложна, поэтому необходимо выполнить процедуру выделения ключевых параметров (feature selection). Признаки были разбиты на количественные и

категориальные. Для количественных признаков были построены графики функций плотности распределения в зависимости от класса, а для категориальных признаков были построены гистограммы. Также был использован коэффициент корреляции Кэнделла для определения степени зависимости параметров от принадлежности классу «отток» или «не отток».

Критерий Кэнделла в данном случае имеет преимущество над критерием корреляции Пирсона, поскольку предназначен для определения зависимости между категориальными, а не количественными признаками. Также критерий Кэнделла является непараметрическим и не предъявляет дополнительных требований к распределению значений в выборках.

Итоговая тепловая карта матрицы корреляции Кэнделла представлена на рисунке 1. По данной матрице видна сильная прямая зависимость оттока от наличия блокировки на лицевого счета абонента, а также от количества обращений. При этом о меньшей склонности к оттоку говорят размеры платежей и объемы трафика, что достаточно очевидно. Из менее явных зависимостей можно выделить обратную зависимость оттока от наличия других услуг, оборудования в рассрочку и возраста.

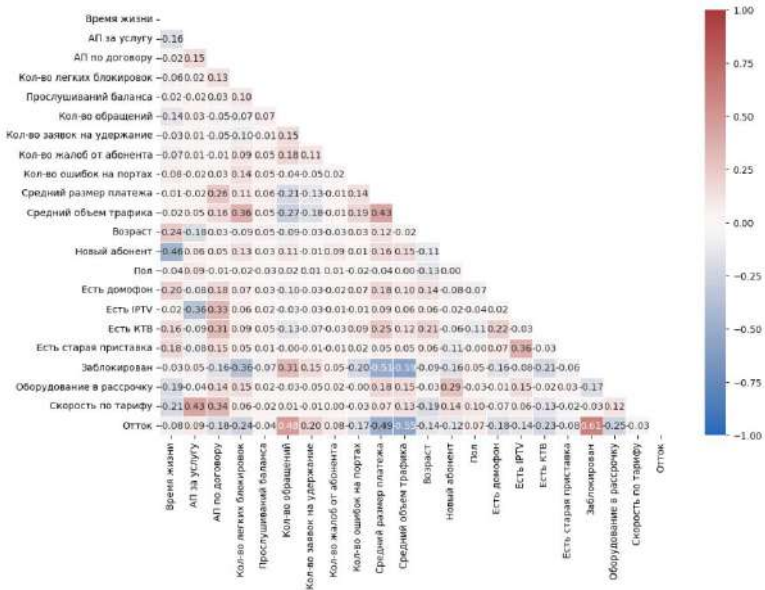


Рис.1 – Тепловая карта корреляционной матрицы датасета

Полученный датасет был разбит на обучающий и тестовый в соотношении 70:30. Результаты тестирования моделей показаны в таблице 1.

**Таблица 1 – Результаты тестирования моделей**

<b>Классификатор</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>	<b>F1 score</b>
Naive Bayes	0.72357	0.66391	0.89745	0.76321
Logistic Regression	0.86113	0.89066	0.82104	0.85443
SVM	0.79932	0.83013	0.74901	0.78749
Decision Tree	0.84761	0.84973	0.8419	0.8458
Random Forest	0.85796	0.94063	0.76196	0.84193

Основной метрикой для оценки качества обученной модели будем считать полноту (Recall), поскольку её максимизация обеспечит выявление большинства клиентов потенциальных для оттока.

По результатам можно сделать вывод, что наибольшую полноту достигает модель наивного Байесовского классификатора, однако она имеет при этом низкий Precision и Accuracy. Более стабильной моделью с максимальной оценкой F1 является дерево решений. Вероятно, модель «случайного леса» на практике должна показать ещё большую эффективность, чем дерево решений, результаты которой будут использованы в дальнейшей работе.

Столь высокие результаты точности предсказаний связаны с достаточно очевидными результатами предсказаний, выражающихся в, например, отсутствии трафика и блокировке лицевого счета абонента. Чаще всего это говорит о том, что абонент не пользуется услугой и очень вероятно расторгнет договор с компанией. Таких абонентов можно выделить без применения методов машинного обучения. Однако рассмотрение не столь очевидных случаев позволит заранее выявить склонность к оттоку и удержать абонентов.

В результате работы был составлен датасет абонентов интернет-провайдера, на основе которого были обучены несколько моделей машинного обучения. Дерево решений показало лучшие результаты метрик при тестировании со значениями Accuracy, Precision, Recall и F1 равными 0,84. Дальнейшие этапы данной работы будут заключаться в отладке описанных ранее моделей, тестировании новых классификаторов, а также проактивной работе со склонными к оттоку абонентами для определения влияния разработанной системы на коэффициент удержания (retention rate) абонентов.

#### Литература:

1. Vafeiadis T. et al. A comparison of machine learning techniques for customer churn prediction // Simulation Modelling Practice and Theory. – 2015. – Т. 55. – С. 1-9.
2. Edwine N. et al. Detecting the risk of customer churn in telecom sector: a comparative study // Mathematical Problems in Engineering. – 2022. – Т. 2022.
3. Bachan L., Gaber T. Predicting customer churn in the internet service provider industry of developing nations: A single, explanatory case study of trinidad and tobago // Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2021. – Springer International Publishing, 2021. – С. 835-844.
4. Radosavljevik D., van der Putten P., Larsen K. K. The impact of experimental setup in prepaid churn prediction for mobile telecommunications: What to predict, for whom and does the customer experience matter? // Trans. Mach. Learn. Data Min. – 2010. – Т. 3. – №. 2. – С. 80-99.
5. Арзамасцев С. А. и др. Предсказание оттока абонентов: сравнение методов машинного обучения // Компьютерные инструменты в образовании. – 2018. – №. 5. – С. 5-23.
6. Карякина А. А., Мельников А. В. Сравнение моделей прогнозирования оттока клиентов интернет-провайдеров // Машинное обучение и анализ данных. – 2017. – Т. 3. – №. 4. – С. 250-256.

## **ПОСТРОЕНИЕ ПРОСТРАНСТВЕННЫХ ИЗОЛИНИЙ ДЛЯ НЕРЕГУЛЯРНЫХ СЕТЕЙ МОНИТОРИНГА**

**Э.Ф. Фарваев**

**Уфимский университет науки и технологий,**

**г. Уфа, farvaev.emil@gmail.com**

**Научный руководитель: Воробьева Г.Р., д.т.н, доцент**

*В данной статье представлены результаты работы по построению пространственных изолиний для визуализации данных на глобусе. Используемый датасет состоит из измерений параметров магнитного поля планеты в местах расположения вариационных станций, в связи с чем точки измерений представляют собой нерегулярную сетку. Также в работе представлен способ избежания разрыва изолиний на линиях географической долготы.*

*This article presents the results of work on constructing spatial isolines for visualizing data on a globe. The dataset used consists of measurements of the*

*parameters of the planet's magnetic field at the locations of variation stations, and therefore the measurement points form an irregular grid. The article also presents a way to avoid breaking isolines on lines of geographic longitude.*

**Введение.** Пространственной изолинией называют линию на карте, в каждой точке которой измеряемая величина имеет одинаковое значение. Изолинии используются в геоинформационных системах для визуализации, анализа, прогноза различных природных и иных явлений. Всем знакомым примером могут служить изолинии высоты над уровнем моря на картах в учебниках по географии. Такие изолинии отражают рельеф поверхности некоторой местности, по ним можно получить представление о наличии холмов, гор, оврагов и т.д.

Одна из основных задач при построении изолиний состоит в необходимости интерполировать известные (как правило, полученные эмпирическим путём) значения в определённых точках для того, чтобы предсказать значения интересующей величины в других точках. Зачастую известные значения распределены по карте нерегулярно по объективным причинам (например, расположение метеостанций). Данный факт обуславливает актуальность построения изолиний при помощи методов, поддерживающих работу с нерегулярными сетями мониторинга.

**1. Постановка задачи.** Целью работы являлось разработать прототип веб API по построению пространственных изолиний. API должен принимать запрос с единственным параметром – временной меткой, в ответ на который возвращать информацию о построенных изолиниях, отражающих уровни определённой величины на всей области земного шара. В качестве формата ответа возьмём GeoJSON [1], являющийся одним из наиболее распространённых среди форматов описания фигур в географических координатах. В качестве визуализируемой изолиниями величины возьмём отклонение географического северного полюса от магнитного. Стоит отметить, что выбор данной величины роли не играет, нам лишь важно, чтобы используемые датасеты содержали замеры нерегулярного характера.

**1.1. Описание исходных данных.** В качестве примера в работе использованы данные о распределении значений геофизических параметров по поверхности Земли. Они получены с ресурса SuperMAG (<https://supermag.jhuapl.edu/>) и представляют собой показания измерений геомагнитного поля, берущиеся с магнитных обсерваторий и вариационных станций [2, 3].

Вышеуказанные данные характеризуются выраженной неоднородностью в виду того, что обсерватории и станции, на которых осуществляются измерения, в основном сосредоточены на территории



Европы и Америки.

Все записи в датасете имеют временную метку и географические координаты места проведения измерения.

**2. Методы интерполяции.** Для обеспечения возможности построения изолиний для определённой величины на некой поверхности или её участке, требуется иметь возможность получить значение данной величины в любой точке обозначенной поверхности/участка. Следовательно, одним из этапов достижения поставленной цели являлась интерполяция имеющихся значений на всю интересующую область исследования.

Методы интерполяции делятся на детерминированные и геостатистические. Детерминированные присваивают значение величины в интерполируемой точке на основе измеренных значений в окрестности этой точки и на заданных математических формулах. Геостатистические основаны на статистических моделях и анализе автокорреляции измеренных точек [4].

**3. Построение изолиний.** Принцип построения изолиний состоит в том, что берётся регулярная сетка со значениями в каждой ячейке, и для каждого значения, которое необходимо визуализировать в виде изолинии, определяется, какие рёбра данная изолиния пересекает, т.е. между какими из ячеек сетки она проходит. Определяются также точки на данных рёбрах, в которых значение величины равно значению изолинии. Эти точки находятся, как правило, при помощи линейной интерполяции значений величины на ребре между ячейками. Далее через найденные точки проводится линия и, при желании, сглаживается [5, 6].

Для прототипа разрабатываемого API выберем метод Clough Tocher, реализуемый методом `interpolate.griddata` библиотеки `scipy` для языка программирования Python [7]. Это детерминированный метод, основанный на подходе интерполяции сплайнами.

Построенные на стороне сервера изолинии будет необходимо передать по сети в виде, удобном для пользователей API, которыми могут являться, например, приложения для визуализации пространственных величин. Своего рода стандартом среди форматов для работы с пространственными данными является GeoJSON. Данный формат основан на JSON и представляет собой открытый стандарт представления пространственных фигур. В данном формате можно передавать точки, линии и полигоны на поверхности земного шара, а также вспомогательные элементы. В качестве системы координат используются геодезические широта и долгота.

В разработанном в процессе работы прототипе построенные по

запросу клиента изолинии трансформируются в GeoJSON и передаются клиенту для дальнейшего использования.

**4. Практическая реализация.** В качестве инструментов воспользуемся библиотеками `numpy`, `scipy`, `matplotlib` и `geojsoncontour` для Python, т.к. они предоставляют богатые возможности для работы с данными, в т.ч. для их интерполяции и визуализации.

Алгоритм работы программного прототипа следующий:

1. Выбрать записи из датасета с указанной в параметре запроса временной меткой и разбить их на два массива. Первый содержит пары чисел: широту и долготу каждой известной точки. Второй содержит значения интересующей величины, в нашем случае – магнитного отклонения;

2. Создать регулярную сетку  $[-90..90] \times [0..360]$  с произвольным шагом. В нашем случае выберем шаг 1. На неё попадут интерполированные значения, по которым далее будут построены изолинии;

3. Интерполировать значения на регулярную сетку при помощи функции `scipy.interpolation.griddata`;

4. По регулярной сетке построить визуализацию изолиниями при помощи библиотеки `matplotlib`;

5. Полученную визуализацию конвертировать в GeoJSON для дальнейшего использования при помощи библиотеки `geojsoncontour`.

**5. Результаты.** Обратим внимание на некоторые особенности разработанного прототипа.

1. Для более наглядных изолиний программа осуществляет масштабирование изолиний. В случаях, когда значения в определённых областях значительно отличаются от остальных, в данных на сетке могут возникать пики. При визуализации изолиниями такие области заметны тем, что в них линии расположены очень близко друг к другу, а в остальных частях сетки они располагаются далеко, что приводит к большим пустым участкам вовсе без линий. Для облегчения анализа тенденций в данных было использовано извлечение квадратного корня из значений для того, чтобы сгладить результирующую поверхность.

2. Многие методы интерполяции работают только в пределах выпуклой оболочки, сформированной точками с известными значениями. Для того, чтобы в результирующих изолиниях не наблюдалось разрывов на одной из линий долготы, программа работает с расширенной регулярной сеткой, т.е. дублирует выборку точек вправо и влево от сетки, на долготы от 360 до 720 и от -360 до 0. Это позволяет увеличить выпуклую оболочку и добиться того, чтобы изолинии не разрывались в месте соединения начала и конца сетки (в нашем случае

это линии долготы 0 и 360 градусов). В случае дальнейшего развития прототипа предлагается реализовать менее ресурсоёмкий способ экстраполяции данных за пределы оболочки.

**Заключение.** В ходе работы был проведён обзорный анализ алгоритма построения изолиний, а также разработан прототип API, предоставляющий данные о построенных изолиниях в формате GeoJSON. Прототип работает с данными нерегулярных сетей мониторинга и осуществляет интерполяцию данных на регулярную сетку для построения изолиний. В процессе разработки были выявлены особенности подобных приложений, в частности необходимость помимо интерполяции данных проводить её экстраполяцию за пределы целевой области, для того, чтобы получить более точные изолинии. Экстраполяция также может помочь в достижении непрерывности изолиний в случае замыкающейся поверхности, например, поверхности глобуса.

#### Литература:

1. The GeoJSON Format. [Электронный ресурс] — URL: <https://www.rfc-editor.org/rfc/rfc7946>.

2. *Gjerloev, J. W.* A Global Ground-Based Magnetometer Initiative / J.W. Gjerloev // *Eos*. — 2009. — Vol. 90, No. 1. — с. 230-231.

3. *Gjerloev, J. W.* The SuperMAG data processing technique / J.W. Gjerloev // *J. Geophys. Res.* — 2012. — No. 117. — с. A09213.

4. Введение в методы интерполяции — ArcGIS Pro | Документация. [Электронный ресурс]. — Режим доступа: <https://pro.arcgis.com/ru/pro-app/latest/help/analysis/geostatistical-analyst/an-introduction-to-interpolation-methods.htm>.

5. *Курахтенков Л. В.* Разработка алгоритма построения изолиний по рассчитанным данным на сферической поверхности / Курахтенков Л. В., Кучумов А. А., Шкиль К. Ю. // *T-Comm – Телекоммуникации и Транспорт*. — 2017. — №5/2017 — с. 21-25.

6. Как работает инструмент Построение изолиний — ArcGIS Pro | Документация. [Электронный ресурс] — URL: <https://pro.arcgis.com/ru/pro-app/latest/tool-reference/3d-analyst/how-contouring-works.htm>.

7. `scipy.interpolate.griddata` — SciPy v1.10.1 Manual. [Электронный ресурс] — URL: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.griddata.html>.

# ИССЛЕДОВАНИЕ ТОЧНОСТИ ОЦЕНКИ ВЕКТОРА СОСТОЯНИЙ НА ОСНОВЕ СИГМА-ТОЧЕЧНОГО ФИЛЬТРА И ЕГО МОДИФИКАЦИИ ПРИ НАЛИЧИИ ВЫБРОСОВ В ИЗМЕРИТЕЛЬНЫХ ДАННЫХ

**В.А. Чернозёмова**

**Новосибирский государственный технический университет,  
г. Новосибирск, vladlena.chernozemova@mail.ru  
Научный руководитель: Черникова О.С., к.т.н., доцент**

*В данной работе приводится сравнение результатов применения сигма-точечного фильтра Калмана и его адаптивной модификации для оценивания вектора состояния при наличии группированных и случайных выбросов в измерительных данных.*

*This paper presents the results of using the unscented Kalman filter and its adaptive modification to estimate the state vector in the presence of grouped and random outliers in the measurement data.*

В настоящее время использование адаптивных фильтров, таких как адаптивный сигма-точечный фильтр Калмана [1], позволяет повысить точность нахождения вектора состояния и, как следствие, точность решения задачи параметрической идентификации динамических систем.

Точность применения алгоритмов калмановской фильтрации при решении практических задач зависит от определения статистических параметров шумов, информация о распределении которых известна неточно или вовсе отсутствует. Применение адаптивных алгоритмов фильтрации позволяет избежать некорректного априорного задания параметров распределения шумов и оценить их на основе измерительных данных [2-4].

В работе исследована точность применения сигма-точечного фильтра Калмана (UKF) и его адаптивной модификации (AUKF) [5] для нахождения оценок вектора состояния стохастической нелинейной непрерывно-дискретной системы при наличии группированных и случайных выбросов в измерительных данных.

Рассмотрим следующую нелинейную модель непрерывно-дискретной системы:

$$\frac{d}{dt}x(t) = -0.1 \cdot \sin(x(t)) + w(t), t \in [0.01, 1], \quad (1)$$

$$y(t_{k+1}) = 0.1 \cdot \sin(2x(t_{k+1})) + v(t_{k+1}), k = 0, 1, \dots, 99. \quad (2)$$

Априорные предположения:

$$x(0) \in N(0, 0.1), w(t) \in N(0, Q(t)), v(t_{k+1}) \in N(0, R(t_{k+1})).$$

О качестве идентификации будем судить по значениям коэффициентов  $\delta_y$ , вычисляющихся по формуле:

$$\delta_y = \frac{\|y_{cp}^* - \hat{y}_{cp}\|}{\|y_{cp}^*\|} = \sqrt{\frac{\sum_{k=0}^{N-1} (y_{cp}^*(t_{k+1}) - \hat{y}_{cp}(t_{k+1}))^2}{\sum_{k=0}^{N-1} (y_{cp}^*(t_{k+1}))^2}},$$

где  $y_{cp}^*(t_{k+1})$  – усредненный незашумленный вектор измерения;  $\hat{y}_{cp}(t_{k+1})$  – усредненный вектор оценки измерений, найденных на основе UKF или AUKF.

Смоделируем выборку измерений, полагая, значения  $Q(t_k) = R(t_{k+1}) = 0.1$  при всех  $k \notin \Omega_A = \{9,10,11,12,13,32,33,34,35,36\}$ , При  $k \in \Omega_A$  увеличим значения ковариационных матриц, моделируя  $\mu = 10$  выбросов.

Полученные численные результаты оценки точности нахождения вектора состояний представлены в следующих табл. 1.

**Таблица 1 – Точность нахождения вектора состояний при наличии группированных выбросов в измерительных данных ( $\mu = 10$ )**

Значения ковариационных матриц шумов при $k \in \Omega_A$	$\delta_y, UKF$	$\delta_y, AUKF$
$Q(t_k) = 1$ $R(t_{k+1}) = 0.5$	0.5308	0.2293
$Q(t_k) = 1.5$ $R(t_{k+1}) = 0.5$	0.5552	0.2899
$Q(t_k) = 0.5$ $R(t_{k+1}) = 0.5$	0.5312	0.2067

Приведем в табл.2 численные результаты точности нахождения вектора состояний при случайном расположении выбросов в измерительных данных, в этом случае  $\Omega_A = \{16,31,40,63,65,70,81,83,89,95\}$ .

Увеличим количество выбросов, при этом аномальные измерения с группированными выбросами расположены в точках:  $\Omega_A = \{10,11,12,13,14,42,43,44,45,46,79,80,81,82,83,91,92,93,94,95\}$ , а случайные –  $\Omega_A = \{1,15,16,18,28,34,36,50,54,56,58,60,63,67,73,76,81,86,88,96\}$ .

Результаты исследований представим в табл. 3, 4.

**Таблица 2 – Точность нахождения вектора состояний при наличии случайных выбросов в измерительных данных ( $\mu = 10$ )**

Значения ковариационных матриц шумов при $k \in \Omega_A$	$\delta_y, UKF$	$\delta_y, AUKF$
$Q(t_k) = 1$ $R(t_{k+1}) = 0.5$	0.3399	0.1868
$Q(t_k) = 1.5$ $R(t_{k+1}) = 0.5$	0.3567	0.2043
$Q(t_k) = 0.5$ $R(t_{k+1}) = 0.5$	0.3461	0.2197

**Таблица 3 – Точность нахождения вектора состояний при наличии группированных выбросов в измерительных данных ( $\mu = 20$ )**

Значения ковариационных матриц шумов при $k \in \Omega_A$	$\delta_y, UKF$	$\delta_y, AUKF$
$Q(t_k) = 1$ $R(t_{k+1}) = 0.5$	0.9078	0.4792
$Q(t_k) = 1.5$ $R(t_{k+1}) = 0.5$	0.8932	0.5878
$Q(t_k) = 0.5$ $R(t_{k+1}) = 0.5$	0.7257	0.5423

**Таблица 4 – Точность нахождения вектора состояний при наличии случайных выбросов в измерительных данных ( $\mu = 20$ )**

Значения ковариационных матриц шумов при $k \in \Omega_A$	$\delta_y, UKF$	$\delta_y, AUKF$
$Q(t_k) = 1$ $R(t_{k+1}) = 0.5$	0.3555	0.2042
$Q(t_k) = 1.5$ $R(t_{k+1}) = 0.5$	0.3088	0.2789
$Q(t_k) = 0.5$ $R(t_{k+1}) = 0.5$	0.3053	0.2366

Экспериментальные результаты показывают, что оценка вектора состояния в пространстве откликов при наличии случайных выбросов более точная, чем при наличии группированных выбросов.

Использование AUKF является целесообразным и предпочтительным при нахождении оценок вектора состояний как при наличии случайных, так и группированных выбросов в измерительных данных и рекомендуется к применению в алгоритмах параметрической идентификации.

Литература:

1. Xia Q. Adaptive fading Kalman filter with an application / Q. Xia, M. Rao, Y. Ying, X. Shen // *Automatica*, – 1994. – V. 30. – № 8. – P. 1333-1338.
2. Mohamed A.H. Adaptive Kalman filtering for INS/GPS / A.H. Mohamed, K.P. Schwarz // *Journal of Geodesy*, – 1999. – V. 73. – P. 193-203.
3. Zheng B. A Robust Adaptive Unscented Kalman Filter for Nonlinear Estimation with Uncertain Noise Covariance / B. Zheng, P. Fu, B. Li, X. Yuan // *Sensors*, 2018. – Vol. 18(3). – 15 p.
4. Gao W. Adaptive Kalman filtering with recursive noise estimator for integrated SINS/DVL systems / W. Gao, J. Li, J. Zhou, Q. Li // *The journal of navigation*, – 2015. – V. 68. – P. 140-161.
5. Chernikova O. S. Two-Stage Parametric Identification Procedure for a Satellite Motion Model Based on Adaptive Unscented Kalman Filters / O. S. Chernikova, A. K. Grechkosev, I. G. Danchenko. – DOI 10.14529/mmp220403. – Text : direct // *Вестник Южно-Уральского государственного университета. Серия: Математическое моделирование и программирование*. – 2022. – Т. 15, № 4. – С. 32–43.

## СОДЕРЖАНИЕ

### НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ

#### Секция ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ПРОЦЕССОВ В ТЕХНОЛОГИЯХ И ПРИРОДНЫХ ЯВЛЕНИЯХ

Авдеева С.Д.	Новосибирск	3
Зотов А.А.	Новосибирск	8
Кислин А.В.	Новосибирск	12
Кислицын С.А., Баранов И.И., Шадрин И.Д.	Нижний Тагил	16
Леонович Д.А.	Новосибирск	20
Продан Д.В.	Москва	25
Ревун А.Л., Рудин С.А., Павский К.В.	Новосибирск	29
Сивенкова А.П.	Новосибирск	33
Скрынник Н.А.	Новосибирск	36
Слобода В.С., Персова М.Г.	Новосибирск	40
Столярова Л.А.	Новосибирск	44
Ульев И.А., Кашкин В.Д.	Новосибирск	48

#### Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО МОДЕЛИРОВАНИЯ

Карасюк А.Д.	Новосибирск	52
Карманова А.А., Сапина Е.А.	Новосибирск	56
Кривецкий А.А.	Новосибирск	62
Кузнецов И.В.	Хабаровск	66
Курочкин О.В.	Новосибирск	71
Никишин А.П.	Санкт-Петербург	74
Патрушев Н.В.	Новосибирск	78
Попов И.М.	Новосибирск	82
Хуснутдинов А. О., Карманов В.С.	Новосибирск	86
Шеховцов Е.С.	Новосибирск	91



**Секция ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ,  
СТАТИСТИКА И ЭКОНОМЕТРИКА**

Альков С.С.	Новосибирск	95
Бортников Н.Е.	Новосибирск	99
Дворецкая В.К.	Новосибирск	103
Кутузова И.А.	Новосибирск	107
Кутырев И.О.	Новосибирск	112
Махин А.А.	Новосибирск	116
Рыбачёк Д.А.	Новосибирск	119
Соколов В.Е.	Санкт-Петербург	122
Табакаева М.Р.	Новосибирск	126
Теников Л.Е., Бодрякова Л.Н.	Новосибирск	130
Чагочкина Д.А.	Новосибирск	134
Шубин А.А.	Новосибирск	138

**Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ,  
АНАЛИЗ И ОБРАБОТКА ДАННЫХ**

Дарбинян А.А., Акопян А.Р.	Ереван	142
Зотов А.А.	Новосибирск	145
Кобылянский В.Г., Кучма И.А.	Новосибирск	149
Легалов Р.А.	Новосибирск	152
Мироненко А.М.	Новосибирск	156
Пекишев Д.В., Коваленко А.В.	Краснодар	161
Полюга В.А., Шабля Ю.В.	Томск	163
Прокудин А.А.	Томск	167
Росуцан А.С.	Москва	171
Рыжков Д.В.	Новосибирск	178
Рябкова М.С., Епанова К.С., Титченко Ю.А.	Нижний Новгород	182
Сухих А.С.	Новосибирск	186
Фарваев Э.Ф.	Уфа	190
Чернозёмова В.А.	Новосибирск	195

**НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ**

**XVII Всероссийская научная конференция молодых ученых  
г. Новосибирск, 04-08 декабря 2023 г.**

**Сборник научных трудов в 11 частях**

**Часть 2**

*Под редакцией Соколовой Д.О.*

Подписано в печать 29.01.2024. Формат 60×84/16. Бумага офсетная.

Уч.-изд. л. 11,85. Печ. л. 12,75. Тираж 100 экз. Заказ № Р-02875

Отпечатано в типографии

Новосибирского государственного технического университета

630073, г. Новосибирск, пр. К. Маркса, 20