ПРИКАСПИЙСКИЙ ЖУРНАЛ

УПРАВЛЕНИЕ И ВЫСОКИЕ ТЕХНОЛОГИИ

2025
№4 (72)

ISSN 2074-1707

16+

ISSN 2074-1707

АСТРАХАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ В. Н. ТАТИЩЕВА

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2025
№ 4 (72)

Журнал включен в перечень рецензируемых научных изданий, рекомендованных ВАК России для публикации основных научных результатов диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук по следующим научным специальностям.

Группа специальностей 1.2 «Компьютерные науки и информатика»:

1.2.2 – Математическое моделирование, численные методы и комплексы программ (технические науки).

Группа специальностей 2.2 «Электроника, фотоника, приборостроение и связь»:

2.2.4 – Приборы и методы измерения (по видам измерений) (технические науки);

2.2.11 – Информационно-измерительные и управляющие системы (технические науки);

2.2.12 – Приборы, системы и изделия медицинского назначения (технические науки).

Группа специальностей 2.3 «Информационные технологии и телекоммуникации»:

2.3.1 – Системный анализ, управление и обработка информации (технические науки);

2.3.4 – Управление в организационных системах (технические науки);

2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки);

2.3.6 – Методы и системы защиты информации, информационная безопасность (технические науки).

Журнал входит в базу данных Ulrich’s Periodicals Directory.

Астрахань
Астраханский государственный университет имени В. Н. Татищева
2025

ББК 30
П75

Рекомендовано к печати редакционно-издательским советом
Астраханского государственного университета имени В. Н. Татищева

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2025
№ 4 (72)

Журнал выходит 4 раза в год
Все материалы, поступающие в редколлегию журнала,
проходят независимое рецензирование

ISSN 2074-1707

ASTRAKHAN TATISHCHEV STATE UNIVERSITY

PRIKASPIYSKIY ZHURNAL:
Upravlenie i Vysokie Tekhnologii

CASPIAN JOURNAL:
Control and High Technologies

A SCIENTIFIC AND TECHNICAL JOURNAL

2025
No. 4 (72)

The journal is included in the list of the reviewed scientific journals recommended by VAK of Russia for the publication of the main scientific results of theses for the candidate of science degree, for the doctor of science degree on the following scientific specialties.

Group of specialties 1.2 "Computer science and informatics":

1.2.2 – Mathematical modelling, numerical methods and complexes of programms (technical sciences).

Group of specialties 2.2 "Electronics, photonics, instrument engineering and communication":

2.2.4 – Instruments and methods of measurement (by type of measurement) (technical sciences);

2.2.11 – Information-measuring and control systems (technical sciences);

2.2.12 – Medical devices, systems and products (technical sciences).

Group of specialties 2.3 "Information technologies and telecommunications":

2.3.1 – System analysis, information control and processing (technical sciences);

2.3.4 – Management in organizational systems (technical sciences);

2.3.5 – Mathematical software and software for computing systems, complexes and computer networks (technical sciences);

2.3.6 – Information security methods and systems, information security (technical sciences).

The journal is included into the database Ulrich's Periodicals Directory.

Astrakhan
Astrakhan Tatishchev State University
2025

Recommended by the Editorial and Publishing Board
of Astrakhan Tatishchev State University

**CASPIAN JOURNAL:
Control and High Technologies**

A SCIENTIFIC AND TECHNICAL JOURNAL

**2025
No. 4 (72)**

The journal is published four times a year
All materials that come to the Editorial Board of the journal
are subject to independent peer-review

СОДЕРЖАНИЕ

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

В. А. Деманов, **И. Ю. Квятковская**, **Г. А. Попов**
Математическая модель и алгоритмы управления
киберфизической системой прудовой аквакультуры 9–14

М. А. Лапина, А. С. Мартынов, К. А. Гедиев, В. Г. Лапин
Методы обнаружения атак нулевого дня
на основе больших языковых моделей 15–23

В. А. Корякова
построение цифрового паттерна поведения пользователя
на основе данных с датчиков смартфона 24–30

Н. В. Ржевская, Д. И. Ледян, М. А. Лапина, В. Г. Лапин
Модель киберрисков цифрового здравоохранения
и регуляторные ограничения 31–41

П. А. Шаронов, А. А. Львов
Оценка методом Монте-Карло неопределенности при построении
аппроксимирующей кривой по экспериментальным данным 42–57

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

И. А. Соловьева, Д. С. Соловьев, С. А. Горохова, А. В. Самохвалов
Выбор оптимальной функции принадлежности для нечетких
множеств на основе экспертной оценки: системный анализ,
математическое моделирование и программная реализация 58–67

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

Р. Ю. Демина, А. В. Хайтул
Модифицированный метод частотного криптоанализа
шифра вертикальной перестановки 68–74

В. А. Частикова, К. В. Козачёк, А. В. Крамарь, Я. М. Маматов
Идентификация ботов в социальных сетях
методами анализа текста и поведения 75–81

М. Н. Жукова, А. О. Яковлева
Решение задачи оптимизации подбора мер
для повышения оценки соответствия требованиям
информационной безопасности кредитных организаций 82–91

В. А. Зюбин, А. С. Макарян, М. М. Путято,
А. Н. Черкасов, М. Е. Косогорова
Разработка программного решения для анализа защищенности
средств безопасного удаленного подключения 92–97

А. С. Вишневский
Сравнительный анализ методов скрытой маркировки
документов на основе стеганографии 98–105

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ МАШИН, КОМПЛЕКСОВ И КОМПЬЮТЕРНЫХ СЕТЕЙ

В. Ю. Кузнецова, Е. Е. Кузнецова, А. В. Попов
Информационная система обучения противодействию атакам методом социальной инженерии 106–114

Р. Э. Асланов, А. А. Большаков, И. В. Вешнева, А. В. Ключиков, Е. И. Малько
Разработка метода проектирования симуляторов виртуальной реальности для подготовки студентов инженерных специальностей на основе онтологической модели 115–125

ПРИБОРОСТРОЕНИЕ, МЕТРОЛОГИЯ И ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ ПРИБОРЫ И СИСТЕМЫ

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ И УПРАВЛЯЮЩИЕ СИСТЕМЫ

Л. Р. Григорьян, Н. М. Богатов, М. С. Коваленко, М. А. Сахно
Построение измерительных каналов в информационно-измерительных системах мониторинга физических процессов 126–132

ПРАВИЛА ДЛЯ АВТОРОВ 133

CONTENTS

INFORMATICS, COMPUTER TECHNIQUE AND CONTROL

SYSTEM ANALYSIS, CONTROL AND INFORMATION PROCESSING

V. A. Demanov, I. Yu. Kvyatkovskaya, G. A. Popov
Mathematical model and control algorithms
of a cyber-physical pond aquaculture system 9–14

M. A. Lapina, A. S. Martynov, K. A. Gediev, V. G. Lapin
Using llm to detect a zero-day attack 15–23

V. A. Koryakova
Building a digital pattern of user behavior
based on data from smartphone sensors 24–30

N. V. Rzhevskaya, D. I. Ledyan, M. A. Lapina, V. G. Lapin
Digital healthcare cyberrisk model
and regulatory restrictions 31–41

P. A. Sharonov, A. A. Lvov
Monte Carlo estimation of uncertainty when constructing
an approximating curve from experimental data 42–57

MATHEMATICAL MODELLING, NUMERICAL METHODS AND PROGRAM SYSTEMS

I. A. Solovjeva, D. S. Solovjev, S. A. Gorokhova, A. V. Samokhvalov
Selection of the optimal membership function for fuzzy sets
based on expert assessment: system analysis, mathematical
modeling and software implementation 58–67

METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY

R. Yu. Demina, A. V. Khaytul
A modified method of frequency cryptanalysis
of a vertical permutation cipher 68–74

V. A. Chastikova, K. V. Kozachek, A. V. Kramar, Yan M., Mamatov
Identification of bots in social networks
by text and behavior analysis methods 75–81

M. N. Zhukova, A. O. Iakovleva
Solving the problem of optimizing merdling selection
to improve compliance with information
security requirements of credit institutions 82–91

**V. A. Zyubin, A. S. Makaryan, M. M. Putyato,
A. N., Cherkasov, M. E. Kosogorova**
Development of a software solution for analyzing
the security of remote secure connection facilities 92–97

A. S. Vishnevsky
Comparative analysis of methods of hidden marking
of digital documents based on steganography 98–105

MATHEMATICAL SOFTWARE AND SOFTWARE FOR COMPUTING MACHINES, COMPLEXES AND COMPUTER NETWORKS

V. Yu. Kuznetsova, E. E. Kuznetsova, A. V. Popov
Information system for training
in countering social engineering attacks 106–114

R. E. Aslanov, A. A. Bolshakov, I. V. Veshneva
A. V. Klyuchikov, E. I. Malko
Development of a method for designing virtual reality simulators
for training engineering students based on an ontological model 115–125

INSTRUMENT ENGINEERING, MEASUREMENT SCIENCE, INFORMATION AND MEASURING DEVICES AND SYSTEMS

INFORMATION, MEASURING AND CONTROL SYSTEMS

L. R. Grigoryan, **N. M. Bogatov**, **M. S. Kovalenko**, **M. A. Sakhno**
Construction of measuring channels in information
and measuring systems for monitoring physical processes 126–132

RULES FOR THE AUTHORS 133

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

УДК 004.021

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И АЛГОРИТМЫ УПРАВЛЕНИЯ КИБЕРФИЗИЧЕСКОЙ СИСТЕМОЙ ПРУДОВОЙ АКВАКУЛЬТУРЫ

Деманов Владимир Анатольевич, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, стр. 16/1,
аспирант, ORCID: 0009-0004-0833-7772, e-mail: vovademanov@mail.ru
Квятковская Ирина Юрьевна, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, стр. 16/1,
доктор технических наук, профессор ORCID: 0000-0001-7205-7231, e-mail: i.kvyatkovskaya @astu.ru
Попов Георгий Александрович, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, стр. 16/1,
доктор технических наук, профессор, e-mail: popov@astu.ru

В работе рассматривается задача построения математической модели и алгоритмов адаптивного управления киберфизической системой прудовой аквакультуры, использующей данные непрерывного мониторинга параметров водной среды. Киберфизическая система трактуется как интегрированный комплекс физических, измерительных, вычислительных и исполнительных подсистем, объединенных замкнутыми контурами обратной связи. Предложена дискретная динамическая модель, описывающая взаимодействие подсистем «среда – объект – вычислительный модуль – исполнительные устройства» с учетом шумов измерений, внешних возмущений и задержек в каналах связи. Разработан алгоритм работы контура «мониторинг – анализ – управление», включающий нормализацию данных, выявление аномалий на основе z-оценок и адаптивное формирование управляющих воздействий по отклонению текущих параметров от заданных. Модель реализована в программном комплексе имитационного моделирования прудового хозяйства. Результаты вычислительных экспериментов показывают, что применение предложенного алгоритма позволяет на 70–80 % уменьшить среднеквадратичное отклонение температуры воды от заданного значения, более чем в 5 раз сократить число выходов контролируемых параметров за допустимые пределы, на 60–70 % уменьшить время пребывания системы в опасных состояниях и снизить энергопотребление исполнительных механизмов на 8–10 % по сравнению с традиционным регламентным управлением.

Ключевые слова: киберфизическая система, прудовая аквакультура, математическая модель, мониторинг, адаптивное управление, имитационное моделирование, качество воды

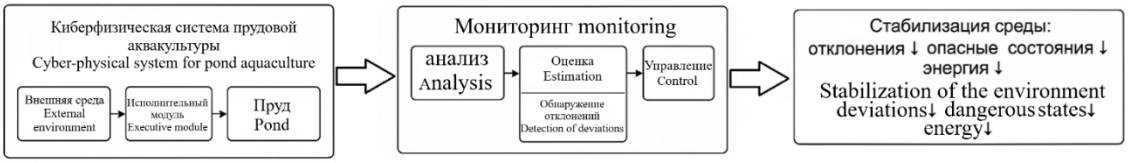
MATHEMATICAL MODEL AND CONTROL ALGORITHMS OF A CYBER-PHYSICAL POND AQUACULTURE SYSTEM

Demanov Vladimir A., Astrakhan State Technical University, bldg 16/1 Tatishev St., Astrakhan, 414056, Russian Federation,
graduate student, ORCID: 0009-0004-0833-7772, e-mail: vovademanov@mail.ru
Kvyatkovskaya Irina Yu., Astrakhan State Technical University, bldg 16/1 Tatishev St., Astrakhan, 414056, Russian Federation,
Doct. Sci. (Engineering), Professor, ORCID: 0000-0001-7205-7231, e-mail: i.kvyatkovskaya@astu.ru
Popov Georgiy A., Astrakhan State Technical University, bldg 16/1 Tatishev St., Astrakhan, 414056, Russian Federation,
Doct. Sci. (Engineering), Professor, e-mail: popov@astu.ru

This article addresses the development of a mathematical model and adaptive control algorithms for a cyber-physical pond aquaculture system that relies on continuous monitoring of water-quality parameters. The cyber-physical system is viewed as an integrated complex of physical, sensing, computing and actuation subsystems coupled by feedback loops. A discrete-time dynamic model is proposed that captures the interaction between the “environment – pond – computing module – actuators” subsystems and accounts for measurement noise, external disturbances and communication delays. An operational algorithm for the “monitoring – analysis – control” loop is designed, incorporating data normalization, anomaly detection based on z-scores and adaptive adjustment of control actions according to deviations of the current parameters from their set points. The model is implemented in a simulation software package for pond aquaculture cyber-physical systems. Numerical experiments demonstrate that the proposed approach reduces the root-mean-square deviation of water temperature from the desired value by 70–80 %, decreases the number of constraint violations by more than a factor of five, shortens the total duration of unsafe states by 60–70 % and lowers the energy consumption of actuators by 8–10 % compared with traditional schedule-based control.

Keywords: cyber-physical system, pond aquaculture, mathematical model, monitoring, adaptive control, simulation modeling, water quality

Graphical annotation (Графическая аннотация)



ВВЕДЕНИЕ

Киберфизические системы (КФС) в аграрном секторе позволяют интегрировать измерительные средства, вычислительные ресурсы и исполнительные механизмы в единый контур управления [2] [3]. Для прудовой аквакультуры это особенно актуально: от поддержания температуры воды, кислотности (pH) и концентрации растворенного кислорода зависит выживаемость и продуктивность выращиваемых гидробионтов [4].

Традиционные системы управления [1] в прудовых хозяйствах базируются на периодическом ручном контроле параметров среды и регламентных действиях (включение аэраторов, подача реагентов). Такой подход плохо учитывает динамику внешних воздействий и временные задержки. В работе предлагается перейти к управлению на основе непрерывного мониторинга и математической модели, описывающей поведение киберфизической системы.

Целью статьи является построение модели киберфизической системы прудовой аквакультуры и разработка алгоритмов управления, использующих данные мониторинга для адаптивного поддержания параметров среды в допустимых диапазонах.

ОБЪЕКТ ИССЛЕДОВАНИЯ И ПОСТАНОВКА ЗАДАЧИ

Объектом исследования является прудовая киберфизическая система, включающая:

– физическую подсистему: водоем с рыбой и гидробионтами, гидротехническими сооружениями;

– подсистему мониторинга: датчики температуры воды, pH, концентрации растворенного кислорода и др.;

– вычислительную подсистему: программно-аппаратный комплекс для сбора и анализа данных;

– подсистему управления: аэраторы, насосы, подогреватели воды, дозаторы реагентов и исполнительные контроллеры [3, 5].

Основная задача управления формулируется как поддержание в пруду параметров среды, которые можно определить формулой (1):

$$X(t) = (T(t),\ pH(t),\ O_2(t),\ \ldots), \tag{1}$$

в заданных диапазонах

$$X_{min} \le X(t) \le X_{max}$$

при минимальных затратах ресурсов (электроэнергии, реагентов, трудозатрат).

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ КИБЕРФИЗИЧЕСКОЙ СИСТЕМЫ

Киберфизическая система описывается в виде динамической модели, включающей состояния внешней среды, физического процесса и управляющих воздействий. В непрерывном времени модель может быть записана в виде системы уравнений (2)–(5):

$$\frac{dE}{dt} = F_E\big(E(t), P(t), U(t)\big) + \Omega_E(t), \tag{2}$$

$$\frac{dP}{dt} = f_P\big(P(t),\ m(t),\ E(t)\big), \tag{3}$$

$$\frac{dC}{dt} = f_C\big(C(t),\ X(t),\ U(t)\big), \tag{4}$$

$$\frac{dU}{dt} = f_U\big(U(t),\ X(t),\ P(t),\ C(t)\big), \tag{5}$$

где E(t) – вектор параметров внешней среды (метеоданные, притоки);

P(t) – вектор параметров водной среды пруда;

C(t) – состояние вычислительной подсистемы (параметры алгоритмов, внутренние переменные); U(t) – вектор управляющих воздействий на исполнительные устройства;

X(t) – измеряемые параметры среды;

m(t) – управляющие воздействия на объект;

$\Omega_E(t)$ – возмущения внешней среды.

Для практической реализации модель удобно представить в дискретном виде с шагом дискретизации Δt, согласованным с периодом опроса датчиков (6)–(7):

$$x(k+1) = f\big(x(k),\ u(k),\ w(k)\big), \tag{6}$$

$$y(k) = h\big(x(k),\, v(k)\big), \tag{7}$$

где k – номер такта;

x(k) – вектор состояний;

u(k) – вектор управляющих воздействий;

y(k) – вектор измеряемых параметров среды;

w(k), v(k) – шумы и возмущения.

В модели выделяются подсистемы:

– физический объект (пруд) (8):

$$x_P(k+1) = f_P\big(x_P(k), u_P(k), v_P(k)\big); \tag{8}$$

– вычислительная подсистема (9):

$$x_C(k+1) = f_C\big(x_C(k), y(k), u_C(k)\big); \tag{9}$$

– канал связи: моделируется задержками и потерями пакетов.

АЛГОРИТМ ФУНКЦИОНИРОВАНИЯ КОНТУРА «МОНИТОРИНГ – АНАЛИЗ – УПРАВЛЕНИЕ»

Алгоритм работы киберфизической системы можно представить как циклический процесс:

1) сбор данных мониторинга;
2) предварительная обработка и анализ;
3) вычисление управляющих воздействий;
4) выполнение управляющих действий;
5) переход к следующему такту [8].

На этапе мониторинга формируется вектор измерений (10):

$$x(k) = \big(T(k),\, pH(k),\, O_2(k)\big). \tag{10}$$

Для повышения качества данных выполняется фильтрация и нормализация. Нормализация осуществляется по формуле (11):

$$x_{norm,i}(k) = \frac{x_i(k) - x_{min,i}}{x_{max,i} - x_{min,i}}, \tag{11}$$

где $x_i(k)$ – текущее значение i-го параметра;

$x_{min,i}$ и $x_{max,i}$ – минимальное и максимальное значения параметра на учебной выборке.

Для диагностики отклонений вводится z-оценка (12):

$$Z_i(k) = \frac{x_i(k) - \mu_i}{\sigma_i}, \tag{12}$$

где μ_i и σ_i – среднее и стандартное отклонение i-го параметра при нормальной работе системы.

На основе $Z_i(k)$ вычисляется вероятность аномального состояния (13):

$$P_{anom}(k) = 1 - \prod_i \big(1 - P_i(k)\big), \tag{13}$$

где $P_i(k)$ – вероятность выхода i-го параметра за допустимые пределы.

Если $P_{anom}(k)$ превышает пороговое значение P_{thr}, формируется сигнал тревоги и запускается процедура корректировки управления.

АЛГОРИТМ АДАПТИВНОГО УПРАВЛЕНИЯ НА ОСНОВЕ ДАННЫХ МОНИТОРИНГА

Управляющие воздействия формируются исходя из отклонения текущих параметров среды от требуемых [7]. Для каждого регулируемого параметра $X_{req,i}$ задается требуемое значение или диапазон. Ошибка управления определяется как (14):

$$E_i(k) = X_{req,i} - X_i(k). \tag{14}$$

На основе вектора ошибок используется алгоритм адаптации управляющих воздействий. В простейшем линейном случае обновление управляющего воздействия $U_i(k)$ можно записать как (15):

$$U_i(k+1) = U_i(k) + \alpha_i\, E_i(k), \tag{15}$$

где α_i – коэффициент адаптации, определяемый чувствительностью объекта.

Для оценки качества управления применяется средняя абсолютная ошибка (MAE), записывающаяся как (16):

$$MAE = \frac{1}{N}\sum_{k=1}^{N} |\, y(k) - \hat{y}(k)\,|, \tag{16}$$

где y(k) – фактические значения параметра,

$\hat{y}(k)$ – заданные (или прогнозируемые) значения;

N – число наблюдений.

В реальной системе алгоритм управления сочетает:

– эмпирические правила (например, включение аэрации при падении кислорода ниже порога) [9];

– адаптивные корректировки коэффициентов по результатам работы;

– прогнозные оценки параметров, получаемые из моделей анализа данных (например, нейросетевых моделей).

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Построенная модель была реализована в программном комплексе имитационного моделирования AnyLogic (рис. 1).

Выбор программного комплекса AnyLogic обусловлен тем, что он поддерживает совместное моделирование непрерывных и дискретных процессов, что принципиально важно для киберфизической системы прудового хозяйства, где сочетаются динамика параметров водной среды и дискретные переключения исполнительных устройств. AnyLogic предоставляет удобные средства задания структурной схемы системы с обратными связями и задержками в каналах связи, визуализации результатов и проведения серий вычислительных экспериментов. Кроме того, данный программный комплекс широко используется в научных и инженерных исследованиях для имитационного моделирования сложных технико-технологических объектов, что повышает воспроизводимость и сопоставимость полученных результатов.

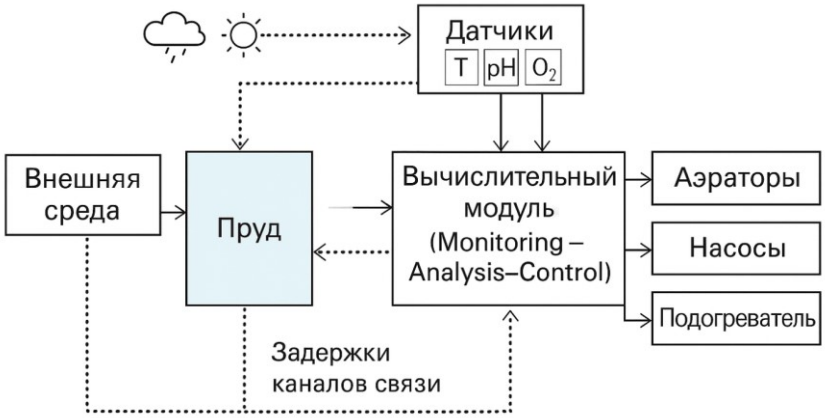


Рисунок 1 – Структурная схема имитационной модели киберфизической системы прудового хозяйства, реализованной в программном комплексе AnyLogic

В модели учитывались суточные колебания температуры и кислорода, изменчивость pH, задержки в каналах связи.

Заполнение модели данными и последующие расчеты показали, что при использовании адаптивного алгоритма управления на основе данных мониторинга удается:

– уменьшить среднеквадратичное отклонение температуры воды от заданной на 70–80 % по сравнению с регламентным управлением;

– сократить число выходов параметров за допустимые пределы более чем в 5 раз;

– снизить суммарное время пребывания системы в «опасных» состояниях примерно на 60–70 %.

Результаты и сравнение с традиционным методом управления показаны на рисунке 2.

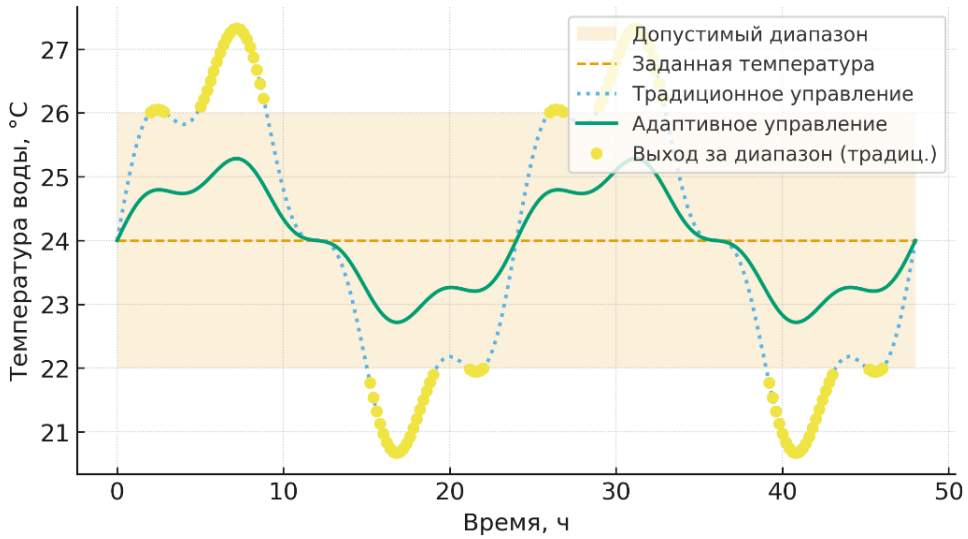


Рисунок 2 – Сравнение изменения температуры воды в пруду при традиционном (красная пунктирная линия) и адаптивном (синяя сплошная линия) управлении

Дополнительно был оценен ресурсный эффект: за счет более рационального включения исполнительных механизмов (аэраторов, насосов) суммарное потребление электроэнергии снизилось на 8–10 % по сравнению с традиционной схемой управления.

ЗАКЛЮЧЕНИЕ

В работе предложена математическая модель киберфизической системы прудовой аквакультуры и алгоритмы управления на основе данных мониторинга. Модель описывает взаимодействие подсистем «среда – объект – вычислительный модуль – исполнительные устройства» и позволяет проводить анализ устойчивости и эффективности различных стратегий управления.

Разработан алгоритм функционирования контура «мониторинг – анализ – управление», включающий нормализацию данных, диагностику аномалий и адаптивное формирование управляющих воздействий. Имитационные эксперименты показали, что использование предложенного подхода позволяет существенно повысить точность поддержания параметров водной среды и снизить ресурсные затраты.

Полученные результаты могут служить основой для дальнейшей разработки интеллектуальных систем управления в аквакультуре и других отраслях, где КФС функционируют в условиях неопределенности и сильных внешних воздействий.

Список источников

1. Борзин, Р. Ю. Управление киберфизическими системами на основе анализа данных мониторинга: отрасли применения, проблемы и методы их решения / Р. Ю. Борзин, А. Г. Кравец // Прикаспийский журнал: управление и высокие технологии. – 2024. – № 1 (65). – С. 5–18.

2. Иванов, И. И. Киберфизические системы в аграрном секторе: состояние и перспективы / И. И. Иванов // Вестник аграрной науки. – 2020. – № 4. – С. 15–24.

3. Кравец, А. Г. Системный анализ и проектирование киберфизических систем / А. Г. Кравец. – Волгоград : ВолгГТУ, 2018. – 210 с.

4. Тевяшов, Г. К. Киберфизические системы в аквакультуре / Г. К. Тевяшов, Р. Ю. Борзин // Цифровизация агропромышленного комплекса : материалы IV Междунар. науч.-практ. конф. – Тамбов, 2024. – С. 112–116.

5. Balakrishnan, S. Design and development of IoT-based smart aquaculture system in a cloud environment / S. Balakrishnan, S. S. Rani, K. C. Ramya // International Journal of Oceans and Oceanography. – 2019. – Vol. 13, № 1. – P. 67–75.

6. Hu, Z. A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture / Z. Hu, Y. Zhang, X. Jiang, H. Wang // Sensors. – 2019. – Vol. 19, № 6. – Article 1420.

7. Liu, X. Adaptive anti-disturbance control of dissolved oxygen in recirculating aquaculture systems / X. Liu, Y. Zhao, Z. Chen, H. Shi // Symmetry. – 2023. – Vol. 15, № 11. – Article 2015.

8. Shete, R. P. IoT-enabled effective real-time water quality monitoring method for aquaculture / R. P. Shete, A. M. Bongale, D. Dharrao // MethodsX. – 2024. – Vol. 13. – Article 102906.

9. Vo, T. T. E. Overview of smart aquaculture system: focusing on applications of machine learning and computer vision / T. T. E. Vo, H. Ko, J.-H. Huh, Y. Kim // Electronics. – 2021. – Vol. 10, № 22. – Article 2882.

References

1. Borzin, R. Yu., Kravets, A. G. Management of cyber-physical systems based on monitoring data analysis: application areas, problems, and solution methods. *Caspian Journal: Control and High Technologies*, 2024, no. 1 (65), pp. 5–18 (In Russ.).

2. Ivanov, I. I. Cyber-physical systems in the agricultural sector: current state and prospects. *Bulletin of Agrarian Science*, 2020, no. 4, pp. 15–24 (In Russ.).

3. Kravets, A. G. *System Analysis and Design of Cyber-Physical Systems.* Volgograd, VolgGTU, 2018. 210 p. (In Russ.).

4. Tevyashov, G. K., Borzin, R. Yu. Cyber-physical systems in aquaculture. *Digitalization of the Agro-Industrial Complex : Proceedings of the IV International Scientific and Practical Conference*, Tambov, 2024, pp. 112–116 (In Russ.).

5. Balakrishnan, S., Rani, S. S., Ramya, K. C. Design and development of an IoT-based smart aquaculture system in a cloud environment. *International Journal of Oceans and Oceanography*, 2019, vol. 13, no. 1, pp. 67–75.

6. Hu, Z., Jiang, X., Wang, H. A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture. *Sensors*, 2019, vol. 19, no. 6, article 1420.

7. Liu, X., Zhao, Y., Chen, Z., Shi, H. Adaptive anti-disturbance control of dissolved oxygen in recirculating aquaculture systems. *Symmetry*, 2023, vol. 15, no. 11, article 2015.

8. Shete, R. P., Bongale, A. M. D. Dharrao IoT-enabled effective real-time water quality monitoring method for aquaculture / R. P. Shete, // MethodsX. – 2024. – vol. 13, article 102906.

9. Vo, T. T. E. Overview of smart aquaculture system: focusing on applications of machine learning and computer vision / T. T. E. Vo, H. Ko, J.-H. Huh, Y. Kim // Electronics, 2021, vol. 10, no. 22, article 2882.

Статья поступила в редакцию 28.11.2025; одобрена после рецензирования 01.12.2025; принята к публикации 05.12.2025.

The article was submitted 28.11.2025; approved after reviewing 01.12.2025; accepted for publication 05.12.2025.

УДК 004.056

МЕТОДЫ ОБНАРУЖЕНИЯ АТАК НУЛЕВОГО ДНЯ НА ОСНОВЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Лапина Мария Анатольевна, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

кандидат физико-математических наук, доцент, ORCID: 0000-0001-8117-9142, e-mail: mlapina@ncfu.ru

Мартынов Андрей Сергеевич, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1; Московский финансово-юридический университет, 117342, Российская Федерация, г. Москва, ул. Введенского, 1а,

аспирант, ORCID: 0009-0007-8349-5481, e-mail: doskam@narod.ru

Гедиев Карим Альбертович, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

бакалавр, ORCID: 0009-0009-6276-8051, e-mail: karimgediev@yandex.ru

Лапин Виталий Геннадьевич, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

кандидат физико-математических наук, доцент, ORCID: 0000-0002-0611-7002, e-mail: vitlx@yandex.ru

В статье рассматривается применение больших языковых моделей (LLM) для решения задачи обнаружения атак нулевого дня. Проведен систематический анализ современных исследований, посвященных возможностям LLM в области кибербезопасности. Особое внимание уделено способности языковых моделей выявлять неизвестные угрозы на основе семантического анализа кода, сетевого трафика и текстовых описаний уязвимостей. Представлен обзор ключевых подходов к интеграции LLM в системы безопасности, включая многоагентные архитектуры, гибридные системы обнаружения вторжений и специализированные фреймворки для распределенных инфраструктур. Проанализированы экспериментальные результаты, демонстрирующие эффективность LLM-решений по сравнению с традиционными методами обнаружения угроз. Подробно исследованы основные вызовы и ограничения, связанные с практическим применением LLM в кибербезопасности. Критически оценены такие проблемы, как уязвимость к состязательным атакам, риск извлечения тренировочных данных, вычислительная сложность и ограничения контекстного окна моделей. В качестве методологии исследования применен систематический обзор литературы. Новизна работы заключается в комплексном анализе возможностей и ограничений LLM именно для обнаружения атак нулевого дня, а также в систематизации перспективных направлений дальнейших исследований в этой междисциплинарной области.

Ключевые слова: большие языковые модели, атаки нулевого дня, кибербезопасность, обнаружение угроз, систематический обзор, машинное обучение, искусственный интеллект, семантический анализ, гибридные системы, MITRE ATT&CK

USING LLM TO DETECT A ZERO-DAY ATTACK

Lapina Maria A., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

Cand. Sci. (Physics and Mathematics), Associate Professor, ORCID: 0000-0001-8117-9142, e-mail: mlapina@ncfu.ru

Martynov Andrey S., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation; Moscow University of Finance and Law, 1a Vvedensky St., Moscow, 117342, Russian Federation,

graduate student, ORCID: 0009-0007-8349-5481, e-mail: doskam@narod.ru

Gediev Karim A., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

bachelor, ORCID: 0009-0009-6276-8051, e-mail: karimgediev@yandex.ru

Lapin Vitaly G., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

Cand. Sci. (Physics and Mathematics), Associate Professor, ORCID: 0000-0002-0611-7002, e-mail: vitlx@yandex.ru

The article discusses the use of large language models (LLM) to solve the problem of detecting zero-day attacks. A systematic analysis of modern research on LLM capabilities in the field of cybersecurity has been conducted. Special attention is paid to the ability of language models to identify unknown threats based on semantic analysis of code, network traffic, and textual descriptions of vulnerabilities. An overview of key approaches to LLM integration into security systems, including multi-agent architectures, hybrid intrusion detection systems, and specialized frameworks for distributed infrastructures, is presented. Experimental results demonstrating the effectiveness of LLM solutions in comparison with traditional threat detection methods are analyzed. The main challenges and limitations associated

with the practical application of LLM in cybersecurity are investigated in detail. Issues such as vulnerability to adversarial attacks, the risk of extracting training data, computational complexity, and limitations of the contextual window of models are critically assessed. A systematic review of the literature was used as the research methodology. The novelty of the work lies in a comprehensive analysis of the capabilities and limitations of LLM specifically for detecting zero-day attacks, as well as in systematizing promising areas of further research in this interdisciplinary field.

Keywords: big language models, zero-day attacks, cybersecurity, threat detection, systematic review, machine learning, artificial intelligence, semantic analysis, hybrid systems, MITRE ATT&CK

ВВЕДЕНИЕ

Атаки нулевого дня (zero-day) представляют серьезную угрозу информационной безопасности. Согласно исследованиям, такая атака использует уязвимость, о которой еще не объявлено публично, что делает защиту практически невозможной до момента раскрытия уязвимости [1]. Статистика показывает, что средняя продолжительность zero-day атаки составляет 312 дней, а после раскрытия уязвимости количество атак увеличивается на 5 порядков [1].

Актуальность проблемы подтверждается современными данными. В 2023 г. было обнаружено 97 уязвимостей нулевого дня, что более чем на 50 % больше, чем в 2022 г. [2]. Наблюдается смещение атак в сторону корпоративных технологий и сторонних компонентов, а также расширение круга групп, использующих такие атаки.

Традиционные методы обнаружения демонстрируют ограниченную эффективность против zero-day угроз. Классические системы обнаружения вторжений сталкиваются с такими проблемами, как семантический разрыв, высокая цена ошибок и неспособность обнаруживать принципиально новые атаки [3]. Машинное обучение лучше справляется с обнаружением вариаций известных атак, чем с новыми угрозами.

Большие языковые модели (Large Language Model, LLM) предлагают новый подход к решению этой проблемы. Исследования демонстрируют, что LLM обладают способностью к мета-обучению и обучению в контексте (in-context learning), что позволяет им адаптироваться к новым задачам без обновления весов модели [4]. Это особенно важно для обнаружения атак нулевого дня, где отсутствуют данные для обучения.

Таксономия MITRE ATT&CK предоставляет фреймворк для применения LLM в кибербезопасности. Модели могут анализировать отчеты и сопоставлять описанное поведение с техниками ATT&CK, генерировать гипотезы для обнаружения и помогать в классификации атак [5]. Это становится критически важным для реагирования на zero-day атаки, когда сигнатуры еще не созданы.

Цель данного обзора – систематизировать современные подходы к применению LLM для обнаружения атак нулевого дня. В работе анализируются возможности LLM в контексте современных вызовов кибербезопасности и рассматриваются перспективы интеграции языковых моделей в системы защиты.

МЕТОДЫ ИССЛЕДОВАНИЯ

Для достижения поставленной цели – систематизации современных подходов к применению больших языковых моделей (LLM) для обнаружения атак нулевого дня – в данном исследовании был применен метод систематического обзора литературы (Systematic Literature Review, SLR). Выбор данной методологии обусловлен необходимостью обеспечения максимальной полноты, объективности и воспроизводимости анализа в условиях быстрого развития новой и междисциплинарной области, находящейся на стыке искусственного интеллекта и кибербезопасности. Исследование проводилось в строгом соответствии с устоявшимися руководствами по проведению SLR в программной инженерии [6].

На первоначальном этапе, этапе планирования, был разработан детальный протокол обзора, который зафиксировал все процедуры и критерии для минимизации систематических ошибок. В рамках протокола были сформулированы ключевые исследовательские вопросы, определившие направление и границы исследования. Вопросы были структурированы с использованием методологии PICOC. В качестве объектов исследования (Population) рассматривались системы и методы обнаружения киберугроз, ориентированные на выявление zero-day атак. Исследуемой интервенцией (Intervention) выступало применение LLM в качестве основного или вспомогательного аналитического компонента. В качестве сравнения (Comparison) рассматривались традиционные методы, включая сигнатурный анализ, анализ аномалий, а также статический и динамический анализ кода. Ключевыми результатами (Outcome) для оценки являлись метрики эффективности, такие как точность (accuracy и precision), полнота (recall), F1-мера, частота ложных срабатываний, а также способность к обобщению на неизвестные угрозы и вычислительная эффективность. Контекст (Context) исследования включал научные работы из областей компьютерной безопасности, программной инженерии и искусственного интеллекта. На основе этой структуры были определены

следующие исследовательские вопросы: какие архитектуры и методы на основе LLM демонстрируют наибольшую эффективность в задачах проактивного обнаружения и прогнозирования zero-day атак; как интеграция LLM с традиционными системами обнаружения влияет на их способность выявлять неизвестные угрозы; каковы ключевые вызовы и ограничения, связанные с применением LLM в данном контексте.

Этап проведения обзора включал систематический поиск, отбор и анализ релевантных научных публикаций. Поиск первоисточников осуществлялся в основных электронных библиографических базах данных, включая IEEE Xplore, ACM Digital Library, SpringerLink, Scopus и Web of Science. Для формирования поисковых строк использовались ключевые термины и их синонимы, такие как "large language model", LLM, GPT, "transformer model", в сочетании с "zero-day", "0-day", "unknown attack", "unknown vulnerability", а также "detection", "cybersecurity" и "intrusion detection". Для минимизации риска предвзятости публикаций и охвата новейших исследований дополнительно проводился ручной поиск в трудах ведущих международных конференций по информационной безопасности, таких как USENIX Security, IEEE Symposium on Security and Privacy и ACM CCS, а также в репозитории препринтов arXiv.

Процесс отбора исследований был многоэтапным и основывался на заранее определенных критериях включения и исключения. Критериями включения являлись: фокус работы на применении LLM для обнаружения или прогнозирования неизвестных уязвимостей или атак; наличие эмпирической оценки предложенных методов; публикация в период с 2020 по 2025 г., отражающая современное состояние области. Из рассмотрения исключались статьи, не прошедшие процедуру рецензирования, а также работы, в которых LLM использовались исключительно для генерации контента, а не для аналитических задач. Процесс отбора включал первоначальный предварительный анализ по заголовкам и аннотациям, последующую детальную оценку по полному тексту и финальное утверждение круга релевантных публикаций.

Для синтеза данных из отобранных исследований применялся концептуально-центричный подход, рекомендованный [7], который предполагает организацию материала вокруг ключевых идей, а не вокруг авторов или публикаций. Из каждой работы извлекались данные об используемой архитектуре LLM, типах анализируемых данных, таких как исходный код, сетевой трафик, логи или текстовые описания уязвимостей, примененных наборах данных для валидации, сравниваемых базовых методах и полученных метриках эффективности. На основе этого была построена концептуальная матрица, которая позволила сгруппировать исследования по основным новым направлениям, таким как многоагентные системы, гибридные системы обнаружения вторжений, специализированные фреймворки и методы проактивного поиска уязвимостей. Данный подход позволил не только систематизировать существующие наработки, но и выявить пробелы в современных исследованиях. Качество каждого включенного в обзор исследования оценивалось по ряду критериев, включая репрезентативность и адекватность использованных наборов данных для задачи обнаружения zero-day атак, наличие корректного сравнения с методами, не основанными на LLM, а также прозрачность методологии. Итогом применения описанной методологии стала структурированная и объективная картина современного состояния области, представленная в последующих разделах данного обзора.

ВОЗМОЖНОСТИ LLM

Большие языковые модели представляют интерес для области кибербезопасности благодаря способности анализировать семантические паттерны. Гипотеза естественности кода предполагает, что программный код обладает статистическими свойствами, аналогичными естественному языку, что открывает возможности для применения языковых моделей в анализе программного обеспечения. Данный подход позволяет идентифицировать аномальные кодовые конструкции, которые могут содержать потенциальные уязвимости или вредоносные фрагменты [8].

Практические исследования демонстрируют различные аспекты применения LLM в безопасности. Анализ работы GitHub Copilot показывает значительные различия в эффективности обнаружения уязвимостей в зависимости от их типа. Модель демонстрирует различную результативность при выявлении межсайтового скриптинга (Cross-Site Scripting, XSS), обхода путей файловой системы (Path Traversal) и внедрения операторов языка структурированных запросов (Structured Query Language, SQL), что указывает на зависимость качества анализа от характера уязвимости [9]. Исследование [10] подтверждает возможность прогнозирования эксплуатации уязвимостей с точностью до 90 % на основе комплексного анализа метаданных и текстовых описаний из специализированных баз данных. Использование 93 578 признаков, включая текстовые поля, метаданные и временные характеристики, обеспечивает высокую достоверность прогнозов.

Опыт промышленного применения LLM в кибербезопасности демонстрирует их практическую ценность. Microsoft сообщает об обработке 65 триллионов сигналов ежедневно с использованием систем на основе искусственного интеллекта [11]. Интеграция языковых моделей в процессы безопасности позволяет автоматизировать анализ угроз, мониторинг аномалий и тестирование

на уязвимости. Особое значение имеет совмещение LLM с таксономией MITRE ATT&CK для категоризации методов поведения злоумышленников. Модели типа GPT-4 показывают эффективность при совместном использовании с традиционными инструментами безопасности, однако их самостоятельное применение для обнаружения атак нулевого дня остается ограниченным [12].

Существенные ограничения сдерживают широкое внедрение LLM в системы безопасности. Исследование GitHub Copilot выявляет, что до 40 % сгенерированного кода содержит уязвимости из списка MITRE Top 25, что создает дополнительный риск при использовании таких систем [9]. К числу критических проблем относятся склонность к генерации недостоверной информации, ограничения контекстного окна и высокая зависимость результатов от формулировок входных запросов [12]. Методология CheckList предлагает систематизированный подход к оценке лингвистических способностей моделей через тестирование базовых компетенций [13].

Для эффективного применения LLM в безопасности необходим комплексный подход, сочетающий языковые модели с традиционными методами анализа и экспертной оценкой [9]. Обучение на актуальных данных и интеграция в операционные процессы безопасности являются обязательными условиями успешного внедрения [11]. Методы поведенческого тестирования рекомендуется применять для верификации устойчивости моделей к различным видам атак, включая промпт-инъекции и манипуляции с входными данными [13]. Перспективным направлением считается развитие специализированных архитектур, адаптированных для задач кибербезопасности, с учетом выявленных ограничений и требований к надежности.

ОБЗОР СОВРЕМЕННЫХ ПОДХОДОВ

Современные исследования в области применения LLM для обнаружения zero-day атак охватывают несколько ключевых направлений. Рассмотрим основные подходы, представленные в научной литературе.

Многоагентные системы демонстрируют существенное преимущество перед одиночными агентами. Исследование [14] представляет систему иерархического планирования и специализированных агентов (Hierarchical Planning and Task-Specific Agents, HPTSA), которая показывает улучшение показателей в 4,3 раза по сравнению с одиночным агентом. Система достигает 42 % успеха на реальных уязвимостях нулевого дня 2024 г. и превосходит по эффективности отдельного в 1,8 раза агента, который заранее знает об уязвимости. Архитектура HPTSA включает иерархического планировщика, менеджера команд и экспертных агентов для конкретных типов уязвимостей (XSS, SQL-инъекции, подделка межсайтовых запросов). Каждый агент имеет специализированные инструменты, документацию по уязвимостям и целевые промпты. Важно отметить, что традиционные сканеры уязвимостей (OWASP ZAP и MetaSploit) показывают 0 % успеха на том же наборе.

Гибридные системы обнаружения вторжений (Intrusion Detection Systems, IDS) сочетают традиционные методы с семантическим анализом LLM. В работе [15] предлагают гибридную систему для сетей интернета вещей (Internet of Things, IoT), которая комбинирует сигнатурный анализ, обнаружение аномалий и семантический анализ GPT-2. Система демонстрирует повышение точности на 6,3 % и снижение ложных срабатываний на 9,0 %. Функция принятия решения реализована как максимум из оценок трех компонентов системы. Эксперименты на наборе данных CSE-CIC-IDS2018 показывают accuracy 98,3 %, precision 97,5 %, recall 96,8 % и F1-меру 97,1 %. При этом система сохраняет производительность, близкую к реальному времени, с задержкой обработки 9,3 мс.

Специализированные фреймворки предлагают комплексные архитектурные решения. Исследование [16] представляет ZeroDay-LLM – фреймворк с гибридной архитектурой «край-центр». На пограничных устройствах используются легковесные энкодеры (двунаправленные энкодерные представления от трансформеров с уменьшением параметров в 78 раз), а центральный движок основан на мощных трансформерных моделях. Система демонстрирует точность 97,8 % на CICIDS2017 и успешно обнаруживает 95,7% неизвестных атак. Частота ложных срабатываний составляет 2,3 %, что на 23 % лучше традиционных систем. Средняя задержка обработки пакета – 12,3 мс, что удовлетворяет требованиям реального времени.

Практические реализации интеграции с существующими инструментами показывают значительное улучшение эффективности. В статье [17] представлен LlamaIDS, который интегрирует традиционные сигнатурные методы Snort с моделью Llama 3.2. После тонкой настройки на структурированных данных из правил Snort и набора данных UNSW-NB15 точность системы возрастает с 43,75до 81,25 %. Частота ложных положительных срабатываний снижается с 42,85 до 14,29 %. Система использует конкретные инструменты: Llama 3.2 (3 миллиарда параметров), Ollama для развертывания и Wireshark для анализа трафика.

Методологии проактивного поиска уязвимостей представляют отдельное перспективное направление. В исследовании [18] предложена система обнаружения атак протоколов с помощью больших языковых моделей (LLM-Assisted Protocol Attack Discovery, LAPRAD) – трехэтапный процесс обнаружения уязвимостей в сетевых протоколах. Методология включает исследование идеи

атаки, генерацию конфигурации атаки и тестирование. С помощью этого подхода обнаружены три новые атаки типа «отказ в обслуживании» на систему доменных имен и переоткрыты две недавно опубликованные атаки. Для обнаружения одной новой атаки требуется от 2 до 7 итераций диалога.

Подходы на основе трансформации сетевых данных демонстрируют высокую эффективность. В работе [19] предложено преобразование сетевых потоков в текстовые предложения вида «Признак1=значение1, Признак2=значение2...». Модель двунаправленных энкодерных представлений от трансформеров в базовой версии без регистрозависимости, обученная только на двух типах атак (распределенный отказ в обслуживании, отказ в обслуживании), показывает точность 99,96 % при тестировании на пяти неизвестных типах атак (зондирование, атака перебором, веб-атака, ботнет, атака с правами пользователя). Метод анализа множественных потоков, объединяющий каждые 4 последовательных потока, позволяет выявлять сложные многоэтапные атаки.

Операционные платформы предлагают системный подход к организации защиты. В статье [20] описан фреймворк для быстрого развертывания защиты, состоящий из трех компонентов. Платформа сбора разведданных об угрозах осуществляет автоматизированный сбор и приоритизацию угроз, используя таксономию десятки главных угроз безопасности веб-приложений для LLM и матрицу тактик и техник противника. Платформа данных агрегирует информацию из телеметрии клиентов и публичных наборов данных. Платформа развертывания использует неизменяемую архитектуру с множественными версиями детекторов.

Специализированные применения в конкретных областях демонстрируют адаптацию подходов к определенным задачам. В исследовании [21] представлен подход обнаружения zero-day атак в интернете транспортных средств с помощью обучения с нулевым количеством примеров (zero-shot) (Zero-Day Bidirectional Encoder Representations from Transformers approach, ZDBERTa). Модель использует оптимизированный метод маскированного языкового моделирования и zero-shot обучение, достигая точности 99,315 % при тестировании на атаках, не встречавшихся при обучении. Для балансировки данных применяется генеративно-состязательная сеть, что значительно улучшает качество модели по сравнению с методами на основе генетических алгоритмов (86,677 % точности).

Эволюция сетевых систем обнаружения вторжений отражает переход к интеграции с LLM. В работе [22] анализируются стратегии адаптации, включающие непрерывное предобучение, контролируемую тонкую настройку и промпт-инжиниринг. Отмечаются такие архитектуры, как BERT/Transformer в качестве энкодера, единые LLM для обнаружения и гибридные подходы. Подчеркивается важность методов параметро-эффективной настройки, таких как низкоранговая адаптация, для снижения вычислительных затрат [22].

Сравнительный анализ представленных подходов показывает разнообразие архитектурных решений и методов применения больших языковых моделей для обнаружения атак нулевого дня. Основные характеристики рассмотренных систем представлены в таблице 1.

Таблица 1 – Сравнительный анализ современных методов

Ссылка	Подход	Метод	Результаты	Область применения
[14]	HPTSA	Многоагентная система	42 % успеха	Веб-безопасность
[15]	Hybrid LLM-IDS	Гибридный анализ	+6,3 % точности	IoT-сети
[16]	ZeroDay-LLM	Архитектура «край-центр»	97,8 % точности	Корпоративные сети
[17]	LlamaIDS	Интеграция с Snort	81,25 % точности	Сетевой трафик
[18]	LAPRAD	Проактивный поиск	3 новых 0-day	Сетевые протоколы
[19]	BERT-SDN	Трансформация данных	99,96 % точности	Программно-конфигурируемые сети
[21]	ZDBERTa	Zero-shot обучение	99,315 % точности	Интернет транспортных средств

ВЫЗОВЫ И ОГРАНИЧЕНИЯ

Применение LLM для обнаружения zero-day атак сталкивается с комплексом фундаментальных ограничений и практических вызовов. Эти ограничения требуют тщательного анализа перед внедрением подобных систем в реальные среды безопасности. Основные категории выявленных ограничений систематизированы в таблице 2.

Таблица 2 – Ключевые ограничения применения LLM

Категория ограничений	Конкретные проблемы	Влияние на обнаружение угроз
Фундаментальные	Эмергентные способности, «стохатический попугай»	Нестабильность результатов, генерация ложных описаний уязвимостей
Безопасностные	Извлечение тренировочных данных, состязательные атаки	Риск утечки данных, возможность обхода детектора
Технические	Ограничение контекстного окна, ресурсоемкость	Невозможность анализа сложных зависимостей, проблемы работы в реальном времени
Функциональные	Низкая надежность анализа кода	Высокий риск пропуска уязвимостей при трансляции кода

Одной из ключевых проблем являются эмергентные способности (emergent abilities) моделей. Исследования показывают, что сложные возможности, такие как логические рассуждения (reasoning) и анализ контекста, проявляются только у моделей значительного масштаба – обычно от 100 миллиардов параметров [23]. Это создает прямую зависимость между эффективностью обнаружения сложных угроз и вычислительными ресурсами, необходимыми для развертывания и эксплуатации таких моделей. При этом решение многих задач, связанных с обнаружением zero-day атак, сопряжено со значительными трудностями. Даже самые современные и крупные модели демонстрируют при их выполнении нестабильные и ненадежные результаты.

Критическим риском безопасности является возможность извлечения тренировочных данных (training data extraction). Атаки, демонстрируемые в исследованиях, позволяют злоумышленникам получать дословные последовательности из обучающего набора модели, включая конфиденциальную информацию об уязвимостях, сигнатуры атак и внутренние данные систем безопасности [24]. Риск извлечения напрямую коррелирует с размером модели: более крупные модели запоминают значительно больше информации. Это создает парадоксальную ситуацию, когда повышение эффективности обнаружения угроз одновременно увеличивает уязвимость самой системы безопасности.

Уязвимость LLM к состязательным атакам (adversarial attacks) также представляет системную угрозу. Состязательные примеры (adversarial examples), специально сконструированные входные данные с малозаметными возмущениями, могут обманывать модели классификации, сохраняя при этом легитимный вид для человеческого восприятия [25]. Данная уязвимость обладает свойствами переносимости между разными моделями и устойчивости к изменениям в тренировочных данных, что делает ее особенно опасной для систем безопасности.

Проблема «стохастического попугая» (stochastic parrot) представляет собой концептуальное ограничение современных больших языковых моделей. Модели комбинируют языковые формы на основе статистических закономерностей, извлеченных из тренировочных данных, без глубокого понимания семантики или причинно-следственных связей [26]. Это приводит к генерации правдоподобных, но фактически неверных описаний уязвимостей и рекомендаций по безопасности, что исключает возможность полного доверия к автономным выводам моделей в критически важных задачах.

Экспериментальные исследования надежности LLM в задачах обработки кода демонстрируют значительные ограничения. При трансляции кода между языками программирования успешные переводы варьируются от 2,1 до 47,3 %, а в реальных проектах эффективность даже современных моделей типа GPT-4 не превышает 8,1 % [27]. Типичные ошибки включают несоответствие поведения интерфейсов программирования приложений (Application Programming Interface, API), удаление критических фрагментов кода, содержащих проверки безопасности, и некорректную обработку входных данных, что может маскировать существующие уязвимости или создавать новые.

Технические ограничения, такие как размер контекстного окна и неспособность анализировать сложные межфайловые зависимости, существенно снижают эффективность LLM для обнаружения атак, использующих нетривиальные взаимодействия в распределенных кодовых базах [27]. Это требует разработки специализированных архитектур и методов анализа, выходящих за рамки стандартных подходов.

Этические и регуляторные вызовы включают риски двойного использования (dual use), необходимость строгого контроля распространения моделей и обеспечения соответствия нормативным требованиям кибербезопасности [28]. Техники взлома моделей (jailbreak), демонстрируемые в исследованиях, показывают возможность обхода встроенных защитных механизмов LLM, что требует разработки дополнительных уровней безопасности и мониторинга.

Ресурсные ограничения, включая значительные энергетические и вычислительные затраты на обучение и эксплуатацию крупных моделей, создают практические барьеры для широкого развертывания систем на основе LLM в организациях с ограниченными бюджетами [26]. Особую

сложность представляют требования к работе в режиме реального времени, характерные для задач обнаружения вторжений.

Для преодоления перечисленных ограничений необходимы комплексные подходы, включающие гибридные архитектуры, сочетающие LLM с традиционными методами статического и динамического анализа, строгий контроль доступа к моделям, тщательное курирование тренировочных данных и регулярный аудит безопасности [24, 25, 29]. Ориентированный на человека дизайн систем позволяет использовать LLM в качестве интеллектуального усилителя (force multiplier) для аналитиков безопасности, а не в качестве полностью автономного решения.

ЗАКЛЮЧЕНИЕ

Проведенный систематический обзор литературы подтверждает перспективность применения LLM для решения сложной задачи обнаружения атак нулевого дня. Анализ современных исследований демонстрирует, что LLM обладают уникальной способностью анализировать семантические паттерны в различных типах данных, включая исходный код, сетевой трафик, системные логи и текстовые описания уязвимостей. Эта способность позволяет выявлять аномалии и потенциальные угрозы без необходимости предварительного обучения на конкретных примерах атак, что особенно важно для противодействия неизвестным угрозам.

Среди рассмотренных подходов наибольшую эффективность показывают многоагентные системы, где специализированные LLM координируют свои действия для поиска уязвимостей, а также гибридные архитектуры, сочетающие языковые модели с традиционными системами обнаружения вторжений. Экспериментальные результаты свидетельствуют, что системы на основе LLM достигают высокой точности обнаружения при тестировании на неизвестных типах атак, в то время как традиционные сигнатурные методы демонстрируют ограниченную эффективность против реальных уязвимостей нулевого дня. Особого внимания заслуживают специализированные фреймворки, разработанные для работы в распределенных инфраструктурах и условиях ограниченных ресурсов.

Несмотря на положительные результаты, практическое применение LLM в системах безопасности сталкивается с существенными ограничениями. К числу основных проблем относятся технические ограничения, связанные с размером контекстного окна моделей и сложностью анализа распределенных кодовых баз. Существенными препятствиями являются уязвимость LLM к состязательным атакам, риск извлечения тренировочных данных и склонность к генерации недостоверной информации. Кроме того, эффективное использование LLM требует значительных вычислительных ресурсов, что может ограничивать их применение в системах реального времени.

Перспективные направления дальнейших исследований включают разработку специализированных архитектур LLM, оптимизированных для задач кибербезопасности, создание стандартизированных наборов данных и методик тестирования, а также исследование методов защиты самих языковых моделей от различных типов атак. Особый интерес представляет развитие гибридных систем, где LLM функционируют как инструмент усиления возможностей аналитиков безопасности, а не как полностью автономные решения. Важным направлением является также адаптация моделей для работы в условиях ограниченных вычислительных ресурсов и обеспечение их интеграции с существующими системами безопасности.

В целом проведенный анализ позволяет сделать вывод о значительном потенциале больших языковых моделей для развития проактивных систем безопасности. Однако для практической реализации этого потенциала необходимы дальнейшие междисциплинарные исследования, объединяющие специалистов в области искусственного интеллекта, кибербезопасности и программной инженерии. Решение существующих проблем и ограничений позволит создать более эффективные системы обнаружения угроз, способные противостоять постоянно развивающимся киберугрозам.

Список источников

1. Bilge, L. Before we knew it: an empirical study of zero-day attacks in the real world / L. Bilge, T. Dumitraş // Proceedings of the 2012 ACM conference on Computer and communications security. – 2012. – P. 833–844.

2. Stone, M. We’re all in this together: A year in review of zero-days exploited in-the-wild in 2023 / M. Stone, J. Semrau, J. Sadowski. – 2024. – URL: https://storage.googleapis.com/gweb-uniblog-publish-prod/documents /YearinReview of ZeroDays.pdf.

3. Sommer, R. Outside the closed world: On using machine learning for network intrusion detection / R. Sommer, V. Paxson // 2010 IEEE symposium on security and privacy. – IEEE, 2010. – P. 305–316.

4. Brown, T. Language models are few-shot learners / T. Brown, et al. // Advances in neural information processing systems. – 2020. – Vol. 33. – P. 1877–1901.

5. Strom, B. E. Mitre att&ck: Design and philosophy: technical report. / B. E. Strom et al. – The MITRE Corporation, 2018.

6. Kitchenham, B. Guidelines for performing systematic literature reviews in software engineering / B. Kitchenham et al. – 2007.

7. Webster, J. Analyzing the past to prepare for the future: Writing a literature review / J. Webster, R. T. Watson // MIS quarterly. – 2002. – P. 13–23.

8. Allamanis, M. A survey of machine learning for big code and naturalness / M. Allamanis et al. // ACM Computing Surveys (CSUR). – 2018. – Vol. 51, № 4. – P. 1–37.

9. Pearce, H. Asleep at the keyboard? Assessing the security of github copilot's code contributions / H. Pearce et al. // Communications of the ACM. – 2025. – Vol. 68, № 2. – P. 96–105.

10. Bozorgi, M. Beyond heuristics: learning to classify vulnerabilities and predict exploits / M. Bozorgi et al. // Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. – 2010. – P. 105–114.

11. Microsoft Digital Defense Report 2023 // Microsoft. – 2023. – URL: https://www.microsoft.com/en-ie/security/security-insider/microsoft-digital-defense-report-2023#section-master-oce07d (дата обращения: 23.11.2025).

12. Achiam, J. Gpt-4 technical report / J. Achiamet al. // arXiv preprint arXiv:2303.08774. – 2023.

13. Ribeiro, M. T. Beyond accuracy: Behavioral testing of NLP models with CheckList / M. T. Ribeiro et al. // arXiv preprint arXiv:2005.04118. – 2020.

14. Zhu, Y. Teams of llm agents can exploit zero-day vulnerabilities / Y. Zhu et al. // arXiv preprint arXiv:2406.01637. – 2024.

15. Al-Hammouri, M. F. Hybrid LLM-Enhanced Intrusion Detection for Zero-Day Threats in IoT Networks / M. F. Al-Hammouri et al. // arXiv preprint arXiv:2507.07413. – 2025.

16. Alsuwaiket, M. A. ZeroDay-LLM: A Large Language Model Framework for Zero-Day Threat Detection in Cybersecurity / M. A. Alsuwaiket // Information. – 2025. – Vol. 16, № 11. – P. 939.

17. Wang, F. LlamaIDS: Real-Time Detection Model of Zero-Day Intrusions Using Large Language Models / F. Wang et al.

18. Aygun, R. LAPRAD: LLM-Assisted PRotocol Attack Discovery / R. Aygun et al. // arXiv preprint arXiv:2510.19264. – 2025.

19. Swileh, M. N. Unseen attack detection in software-defined networking using a BERT-based large language model / M. N. Swileh, S. Zhang // AI. – 2025. – Vol. 6, № 7. – P. 154.

20. Swanda, A. A Framework for Rapidly Developing and Deploying Protection Against Large Language Model Attacks / A. Swanda et al. // arXiv preprint arXiv:2509.20639. – 2025.

21. Mirza, A. ZDBERTa: Advancing Zero-Day Cyberattack Detection in Internet of Vehicle with Zero-Shot Learning / A. Mirza et al. // Computers. – 2025. – Vol. 14, № 10. – P. 424.

22. Feng, Y. Network Intrusion Detection: Evolution from Conventional Approaches to LLM Collaboration and Emerging Risks / Y. Feng, K. Sakurai // arXiv preprint arXiv:2510.23313. – 2025.

23. Wei, J. Emergent abilities of large language models / J. Wei et al. // arXiv preprint arXiv:2206.07682. – 2022.

24. Carlini, N. Extracting training data from large language models / N. Carlini et al. // 30th USENIX security symposium (USENIX Security 21). – 2021. – P. 2633–2650.

25. Szegedy, C. Intriguing properties of neural networks / C. Szegedy et al. // arXiv preprint arXiv:1312.6199. – 2013.

26. Bender, E. M. On the dangers of stochastic parrots: Can language models be too big? / Bender E. M. et al. // Proceedings of the 2021 ACM conference on fairness, accountability, and transparency. – 2021. – P. 610–623.

27. Pan, R. Lost in translation: A study of bugs introduced by large language models while translating code R. Pan et al. // Proceedings of the IEEE/ACM 46th International Conference on Software Engineering. – 2024. – P. 1–13.

28. Ji, J. Ai alignment: A comprehensive survey / J. Ji et al. // arXiv preprint arXiv:2310.19852. – 2023.

29. Chao, P. Jailbreaking black box large language models in twenty queries / P. Chao et al. // 2025 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML). – IEEE, 2025. – P. 23–42.

References

1. Bilge, L., Dumitraş, T. Before we knew it: an empirical study of zero-day attacks in the real world. *Proceedings of the 2012 ACM conference on Computer and communications security*, 2012, pp. 833–844.

2. Stone, M., Semrau, J., Sadowski, J. *We're all in this together: A year in review of zero-days exploited in-the-wild in 2023*, 2024. Available at: https://storage. googleapis.com/gweb-uniblog-publish-prod/documents/YearinReviewofZeroDays.pdf.

3. Sommer, R., Paxson, V. Outside the closed world: On using machine learning for network intrusion detection. *2010 IEEE symposium on security and privacy*. IEEE, 2010, pp. 305–316.

4. Brown, T. et al. Language models are few-shot learners. *Advances in neural information processing systems*, 2020, vol. 33, pp. 1877–1901.

5. Strom, B. E. et al. *Mitre att&ck: Design and philosophy: technical report.* The MITRE Corporation, 2018.

6. Kitchenham, B. et al. *Guidelines for performing systematic literature reviews in software engineering*, 2007.

7. Webster J., Watson R. T. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, 2002, pp. 13–23.

8. Allamanis, M. et al. A survey of machine learning for big code and naturalness. *ACM Computing Surveys (CSUR)*, 2018, vol. 51, no. 4, pp. 1–37.

9. Pearce, H. et al. Asleep at the keyboard? Assessing the security of github copilot's code contributions. *Communications of the ACM*, 2025, vol. 68, no. 2, pp. 96–105.

10. Bozorgi, M. et al. Beyond heuristics: learning to classify vulnerabilities and predict exploits. *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010, pp. 105–114.

11. *Microsoft Digital Defense Report 2023 (2023). Microsoft.* Available at: https://www.microsoft.com/en-ie/security/security-insider/microsoft-digital-defense-report-2023#section-master-oce07d (accessed 23.11.2025).

12. Achiam, J. et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

13. Ribeiro, M. T. et al. Beyond accuracy: Behavioral testing of NLP models with CheckList. *arXiv preprint arXiv:2005.04118*, 2020.

14. Zhu, Y. et al. Teams of llm agents can exploit zero-day vulnerabilities. *arXiv preprint arXiv:2406.01637*, 2024.

15. Al-Hammouri, M. F. et al. Hybrid LLM-Enhanced Intrusion Detection for Zero-Day Threats in IoT Networks. *arXiv preprint arXiv:2507.07413*, 2025.

16. Alsuwaiket, M. A. ZeroDay-LLM: A Large Language Model Framework for Zero-Day Threat Detection in Cybersecurity. *Information*, 2025, vol. 16, no. 11, pp. 939.

17. Wang, F. et al. *LlamaIDS: Real-Time Detection Model of Zero-Day Intrusions Using Large Language Models.*

18. Aygun, R. et al. LAPRAD: LLM-Assisted PRotocol Attack Discovery. *arXiv preprint arXiv:2510.19264*, 2025.

19. Swileh, M. N., Zhang, S. Unseen attack detection in software-defined networking using a BERT-based large language model. *AI*, 2025, vol. 6, no. 7, p. 154.

20. Swanda, A. et al. A Framework for Rapidly Developing and Deploying Protection Against Large Language Model Attacks. *arXiv preprint arXiv:2509.20639*, 2025.

21. Mirza, A. et al. ZDBERTa: Advancing Zero-Day Cyberattack Detection in Internet of Vehicle with Zero-Shot Learning. *Computers*, 2025, vol. 14, no. 10, p. 424.

22. Feng, Y., Sakurai, K. Network Intrusion Detection: Evolution from Conventional Approaches to LLM Collaboration and Emerging Risks. *arXiv preprint arXiv:2510.23313*, 2025.

23. Wei, J. et al. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*, 2022.

24. Carlini, N. et al. Extracting training data from large language models. *30th USENIX security symposium (USENIX Security 21)*, 2021, pp. 2633–2650.

25. Szegedy, C. et al. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*, 2013.

26. Bender, E. M. et al. On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 2021, pp. 610–623.

27. Pan, R. et al. Lost in translation: A study of bugs introduced by large language models while translating code. *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering*, 2024, pp. 1–13.

28. Ji, J. et al. Ai alignment: A comprehensive survey. *arXiv preprint arXiv:2310.19852*, 2023.

29. Chao, P. et al. Jailbreaking black box large language models in twenty queries. *2025 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML).* IEEE, 2025, pp. 23–42.

Статья поступила в редакцию 27.11.2025; одобрена после рецензирования 02.12.2025; принята к публикации 05.12.2025.

The article was submitted 27.11.2025; approved after reviewing 02.12.2025; accepted for publication 05.12.2025.

УДК 303.732

ПОСТРОЕНИЕ ЦИФРОВОГО ПАТТЕРНА ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЯ НА ОСНОВЕ ДАННЫХ С ДАТЧИКОВ СМАРТФОНА

Корякова Виктория Андреевна, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, стр. 16/1; Астраханский государственный университет им. В. Н. Татищева, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

аспирант, ассистент кафедры информационной безопасности АГУ им. В. Н. Татищева, ORCID: 0000-0002-2472-9248, e-mail: koryakova-01@mail.ru

В статье рассматривается проблема формирования достоверного цифрового паттерна поведения пользователя на основе комплексных данных, собираемых встроенными датчиками смартфона. Актуальность исследования обусловлена ростом числа мобильных сервисов, требующих точной идентификации и анализа активности пользователя. Цель работы – разработка методики, позволяющей выделять устойчивые поведенческие характеристики по данным акселерометра, гироскопа. Полученные результаты включают выявление значимых корреляций между типами активности и параметрами данных с датчиков, а также построение поведенческого профиля, демонстрирующего устойчивость при повторных измерениях. Новизна работы заключается в интеграции разнородных данных с датчиков смартфона в единый паттерн с оценкой надежности, что позволяет существенно повысить точность распознавания пользовательских сценариев.

Ключевые слова: поведенческий паттерн, данные с датчиков смартфона, цифровой паттерн пользователя, машинное обучение, искусственный интеллект, мобильная идентификация

BUILDING A DIGITAL PATTERN OF USER BEHAVIOR BASED ON DATA FROM SMARTPHONE SENSORS

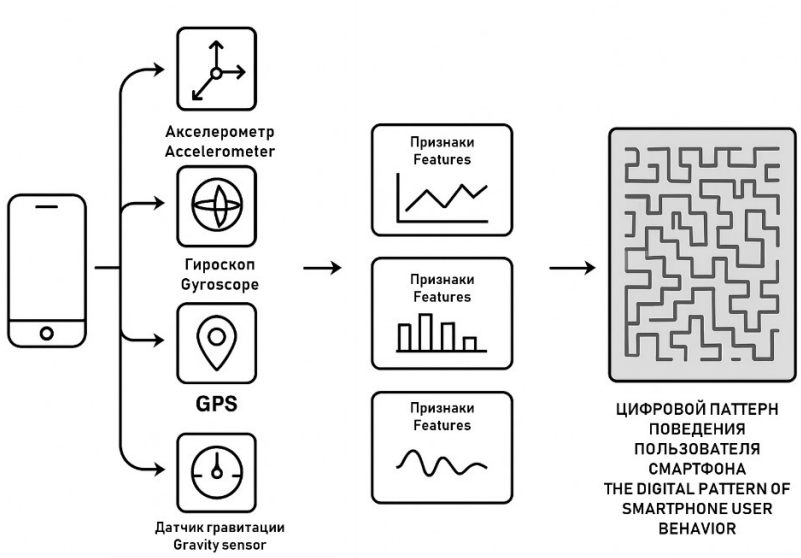
Koryakova Victoria A., Astrakhan State Technical University, bldg. 16/1 Tatishchev St., Astrakhan, 414056, Russian Federation; Astrakhan Tatishchev State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

graduate student, assistant of the Department of Information Security, ORCID: 0000-0002-2472-9248, e-mail: koryakova-01@mail.ru

The article discusses the problem of forming a reliable digital pattern of user behavior based on complex data collected by the smartphone's built-in sensors. The relevance of the study is due to the growing number of mobile services that require accurate identification and analysis of user activity. The aim of the work is to develop a methodology that makes it possible to identify stable behavioral characteristics based on accelerometer and gyroscope data. The results obtained include the identification of significant correlations between activity types and sensor data parameters, as well as the construction of a behavioral profile demonstrating stability during repeated measurements. The novelty of the work lies in the integration of heterogeneous data from smartphone sensors into a single pattern with an assessment of reliability, which significantly improves the accuracy of recognizing user scenarios.

Keywords: behavioral pattern; data from smartphone sensors; digital user pattern; machine learning; artificial intelligence; mobile identification

Graphical annotation (Графическая аннотация)



ВВЕДЕНИЕ

Развитие технологий позволили широко внедрить пространственно-временные датчики в различные приборы, в том числе смартфоны, которые стали неотъемлемой частью повседневной жизни людей. С развитием сектора разработки мобильных приложений современный гаджет получает все больше функциональных возможностей. Пространственно-временные датчики смартфона позволяют собирать данные о местоположении, движении и положении устройства. На основе информации, снимаемой с таких датчиков, можно построить поведенческие паттерны пользователя для дальнейшего использования в процессах аутентификации, прогнозирования состояния человека, в процессах передачи визуальной информации, когда необходимо оцифровать определенные жесты. Совершенствование методов сбора, предобработки и анализа данных создает новый уровень понимания человеческого поведения и его взаимосвязи с окружающей средой. Разработка моделей, методик и алгоритмов для эффективного построения цифрового паттерна поведения человека позволит повысить точность и надежность идентификации типичного и атипичного поведения человека, а также создать новые возможности для применения биометрических систем в различных областях.

Объектом исследования является цифровой паттерн поведения человека. Предметом исследования являются модели, методики и алгоритмы обработки информации с пространственно-временных датчиков смартфона.

Цель исследования – создание методики обработки информации с пространственно-временных датчиков смартфона и построение цифровых поведенческих паттернов человека для дальнейшего использования в процессах аутентификации, передачи данных о состояниях индивидуума.

Научная новизна исследования заключается в разработке решений, направленных на сбор, обработку информации, получаемых от пользователя смартфона; разработке методики для распознавания поведенческих черт пользователя смартфона и методики формирования наборов данных для обучения модели; выработке критериев и метрик для распознавания типичного и атипичного поведения на основе анализа траекторий перемещения мобильного устройства.

ЦИФРОВОЙ ПАТТЕРН ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЯ

Цифровой паттерн поведения пользователя – это совокупность действий, взаимодействий и привычек, которые пользователь демонстрирует при использовании цифровых технологий, таких как веб-сайты, мобильные приложения и социальные сети.

Построение цифрового паттерна поведения пользователя необходимо для решения ряда ключевых задач, связанных с анализом, персонализацией и обеспечением безопасности мобильных систем. Прежде всего, цифровой паттерн позволяет сформировать поведенческий профиль, основанный на объективных данных датчиков смартфона (рис. 1). Такой профиль отражает реальные привычки пользователя, его манеру взаимодействия с устройством, характер перемещения и временную структуру активности [1]. Во-первых, это может повысить точность поведенческой аутентификации, позволяя идентифицировать пользователя не по паролям или биометрии, а по его уникальной моторике и динамике движений. Во-вторых, цифровой паттерн может использоваться в системах персонализации сервисов, адаптируя интерфейсы и реакции приложений к индивидуальным особенностям. В-третьих, он обеспечивает возможность раннего выявления аномалий, возникающих при мошеннических действиях, подмене устройства или нарушении типичных сценариев поведения. Кроме того, цифровой паттерн служит основой для оптимизации мобильных приложений, анализа пользовательской нагрузки и проектирования более удобных интерфейсов [2].

Рисунок 1 – Цифровой паттерн поведения пользователя

Цифровой паттерн включает в себя данные о том, как пользователь перемещается по интерфейсу, какие страницы или функции он использует, сколько времени проводит на них, а также его реакции на различные элементы, такие как кнопки, уведомления или контент. К ключевым характеристикам, влияющим на построение цифрового паттерна, относятся: индивидуальная моторика (манера держать и перемещать смартфон), типичные маршруты, ритмика шагов, особенности взаимодействия с экраном, а также временная структура активности – повторяемость действий в течение суток и недели.

Особенность построения цифрового паттерна состоит в необходимости объединения разнородных потоков данных с датчиков смартфона, различающихся по частоте, шумам и точности. Для этого применяются методы нормализации, фильтрации и синхронизации временных рядов, позволяющие выделить информативные признаки и исключить случайные колебания. Комбинирование данных акселерометра, гироскопа, GPS, датчиков ориентации и барометра создает многомерное представление поведения, устойчивое к внешним помехам и случайным артефактам.

Необходимость формирования такого паттерна обусловлена ростом задач, требующих высокой степени персонализации и надежной поведенческой аутентификации. Цифровой паттерн позволяет автоматизировать идентификацию пользователя, повысить безопасность сервисов, оптимизировать интерфейсы и адаптировать мобильные приложения под индивидуальные особенности поведения, что делает его важным инструментом в современных системах анализа пользовательской активности.

ОБЗОР ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ДАТЧИКОВ

Пространственно-временные датчики смартфона образуют основу для высокоточного построения цифрового паттерна поведения пользователя, фиксируя его перемещения, ориентацию устройства и динамику взаимодействия с окружающей средой. К ключевым датчикам относятся акселерометр, гироскоп, магнитометр, датчик шагов, барометр и GPS-модуль. Акселерометр и гироскоп обеспечивают регистрацию линейных и угловых ускорений, что позволяет выявлять микродвижения, характерные для конкретных моделей использования смартфона. Магнитометр дополняет сведения об ориентации в пространстве, повышая стабильность интерпретации данных. GPS и сетевые методы позиционирования предоставляют точные пространственные координаты и траектории перемещения, а барометр помогает определять изменение высоты, уточняя вертикальную динамику активности.

Акселерометр измеряет ускорение и позволяет смартфону определять характеристики движения и положения в пространстве. Именно этот датчик работает, когда вертикальная ориентация меняется на горизонтальную при повороте устройства. Он же отвечает за подсчет шагов и измерение скорости движения во всевозможных приложениях-картах. Акселерометр дает информацию о том, в какую сторону повернут смартфон. Акселерометр различает движения по трем осям. С его помощью смартфон меняет ориентацию экрана с вертикальной на горизонтальную, позволяет заглушить сигнал вызова, перевернув устройство экраном вниз, а также ориентироваться в картах и считать пройденные шаги. Все конструктивные элементы размещаются в чипе. К неподвижному корпусу на упругих приставках, которые позволяют перемещение в определенных пределах, крепится перегородка с отведенными в сторону проводниками. Эти отводы размещаются между контактами, которые и снимают показания [3]. При перемещении отводов напряженность поля вокруг контактов меняет свои характеристики, что и служит показателем для измерения.

С акселерометром обычно связывают следующие параметры:

– масштабный коэффициент – коэффициент пропорциональности между измеряемым кажущимся ускорением и выходным сигналом (электрическим сигналом, частотой колебаний (для струнного акселерометра) или цифровым кодом);

– пороговая чувствительность (разрешение) – величина минимального изменения кажущегося ускорения, которое способен определить прибор;

– смещение нуля – показания прибора при нулевом кажущемся ускорении;

– случайное блуждание – среднеквадратичное отклонение от смещения нуля;

– нелинейность – изменения зависимости между выходным сигналом и кажущимся ускорением при изменении кажущегося ускорения.

Гироскоп – это устройство, которое может измерять изменения углов ориентации связанного с ним тела относительно инерциальной системы координат. Принцип работы гироскопа заключается в грузиках, которые вибрируют на плоскости с частотой скорости, умноженной на перемещение. При повороте гироскопа возникает так называемое Кориолисово ускорение [4].

В мобильных телефонах гироскоп обычно работает в паре с акселерометром. За счет такого взаимодействия повышается чувствительность гаджета к любому наклону или повороту.

Принцип работы гироскопа основан на законе сохранения углового момента, который гласит, что если на тело не действуют внешние моменты сил, то его угловой момент остается постоянным.

Гироскоп в мобильном устройстве имеет три оси вращения – ось X, ось Y и ось Z. Когда устройство поворачивается или вращается вокруг одной из этих осей, гироскоп регистрирует изменение угловой скорости и передает данные процессору устройства. Этот процесс позволяет устройству определять свою ориентацию в пространстве и реагировать соответствующим образом на движения пользователя. Гироскоп также используется для стабилизации изображения при съемке видео на мобильном устройстве, для управления играми с помощью жестов или вращения устройства, а также для других приложений, где требуется точное определение угловой скорости устройства.

Датчик гравитации в мобильном устройстве работает на основе ускорения свободного падения, которое вызывает гравитационное притяжение Земли [5].

Принцип работы датчика гравитации заключается в измерении ускорения, которое действует на устройство в пространстве. Датчик состоит из устройства, способного регистрировать изменения ускорения в трех основных направлениях: ось X (горизонтальное перемещение влево и вправо), ось Y (горизонтальное движение вперед и назад) и ось Z (вертикальное движение вверх и вниз).

При изменении положения устройства, например, при повороте или наклоне, датчик гравитации регистрирует соответствующие изменения ускорения в каждом из направлений. Эта информация затем передается на устройство, где может быть использована для выполнения различных функций, таких как автоповорот экрана, управление игровыми приложениями или измерение физической активности.

Таким образом, датчик гравитации в мобильном устройстве помогает определять ориентацию и положение устройства в пространстве, что делает его полезным инструментом для удобного взаимодействия с устройством и использования различных приложений и функций.

Сочетание этих датчиков дает возможность формировать многомерное описание поведения пользователя с высокой временной разрешающей способностью. Интеграция данных позволяет отслеживать устойчивые модели движения, распознавать типы активности и строить персонализированные профили, применимые в задачах безопасности, персонализации и анализа повседневных привычек.

ПОВЕДЕНЧЕСКИЕ ХАРАКТЕРИСТИКИ ПРИ ПОСТРОЕНИИ ПАТТЕРНА ПОВЕДЕНИЯ

Типичное поведение пользователя смартфона на основе данных с пространственно-временных датчиков включает в себя регулярные и предсказуемые действия, которые можно наблюдать у большинства пользователей. Например, многие люди используют свои смартфоны в течение дня для проверки уведомлений, общения через мессенджеры или социальных сетей, а также для навигации. Атипичное поведение может проявляться в отклонениях от этих норм. Пользователь может демонстрировать необычные паттерны активности, использование смартфона в условиях, когда это обычно не делается. Для каждого пользователя смартфона присуще уникальное поведение при взаимодействии со смартфоном. Исходя из этого, в рамках исследования планируется разработать методику формирования эталонного поведения, определить метрики типичного и атипичного поведения человека, а также разработать методику обработки информации с пространственно-временных датчиков смартфона для построения цифрового паттерна поведения человека.

Поведенческие характеристики, используемые при построении цифрового паттерна, представляют собой совокупность статистически устойчивых признаков, отражающих индивидуальные особенности взаимодействия пользователя со смартфоном. Они формируются на основании пространственно-временных данных датчиков и описывают как микродинамику движений, так и более крупные поведенческие сценарии. Поведенческие характеристики включают кинематические особенности движения устройства, такие как амплитуда и частота линейных ускорений, угловые скорости вращения, а также характер удержания смартфона и микродинамику движений. Существенную роль играет локомоционная активность, выражающаяся в ритме и длине шага, скорости перемещения и изменении высоты, что фиксируется за счет работы акселерометра, гироскопа, GPS. Пространственные характеристики формируются на основании закономерностей перемещений в течение суток или недели, а временные особенности отражают цикличность действий и повторяемость активности.

Сбор данных осуществляется датчиками смартфона, работающими непрерывно или в режиме периодического опроса (табл.).

Таблица – Сбор данных с датчиков смартфона

Датчик	Назначение
Акселерометр	Моторика, шаги, динамика взаимодействия
Гироскоп	
GPS и сетевое позиционирование	Маршруты, скорость
Барометр	Изменение высоты
Датчики ориентации	Положение устройства в пространстве

Каждая запись снабжается временной меткой, что позволяет выстраивать последовательности событий и выявлять устойчивые поведенческие циклы. После сбора данные проходят предварительную фильтрацию для удаления шумов и артефактов, затем нормализуются и синхронизируются, поскольку разные датчики работают с различной частотой и уровнем точности. На следующем этапе выделяются информативные признаки – интегральные показатели, спектральные характеристики, параметры траектории и статистические метрики, которые затем объединяются в многомерное представление поведения. Сформированный цифровой паттерн проходит проверку на устойчивость и повторяемость, что позволяет использовать его для задач аутентификации, персонализации и анализа активности пользователя.

Типичное поведение пользователя смартфона характеризуется повторяющимися и предсказуемыми паттернами, отражающими его повседневные привычки. К таким признакам относятся постоянные маршруты, например, путь от дома до работы и обратно, а также регулярные точки интереса, такие как магазин, спортзал или кафе. Время пребывания в ключевых локациях обычно стабильно: дома пользователь находится ночью, на работе днем. Его перемещения в течение недели, как правило, предсказуемы. Что касается активности приложений, пользователь регулярно использует определенные программы: мессенджеры утром и вечером, социальные сети в течение дня, карты по необходимости. Время экрана также имеет закономерности – утренние и вечерние периоды активного использования. Данные с датчиков отражают привычные движения, такие как ходьба или поездки на транспорте, а также естественные частоты касаний экрана и типичные жесты. Социальное и коммуникационное поведение включает регулярные звонки и сообщения ограниченному кругу контактов, а также предсказуемые паттерны онлайн-активности, например, проверку почты или социальных сетей в определенные периоды [6, 7].

Атипичное поведение, наоборот, проявляется в отклонениях от привычного паттерна. Это могут быть новые маршруты или точки интереса, не соответствующие привычному распорядку, необычные временные паттерны, например ночные перемещения, частые перемещения между локациями без очевидной цели. В активности приложений проявляются необычные действия: использование новых программ, длительное или непривычное время активности, частое переключение между приложениями без закономерности. Датчики могут фиксировать необычные движения смартфона, резкие падения или длительное отсутствие движения, когда пользователь обычно активен, а также непривычные паттерны касаний экрана, которые могут указывать на автоматизированные действия. Социальное и коммуникационное поведение атипичного пользователя проявляется в резком увеличении числа звонков или сообщений неизвестным контактам, а также в необычных географических или временных совпадениях активности с другими устройствами [8].

МЕТОДИКА ОБРАБОТКИ ИНФОРМАЦИИ С ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ДАТЧИКОВ СМАРТФОНА ДЛЯ ПОСТРОЕНИЯ ЦИФРОВОГО ПАТТЕРНА ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЯ

Методика обработки информации с пространственно-временных датчиков смартфона для построения цифрового паттерна поведения пользователя основывается на последовательной обработке потока данных с датчиков смартфона в устойчивое многомерное представление, отражающее характерные особенности активности (рис. 2, 3).

Первым этапом является сбор данных с акселерометра, гироскопа, датчиков ориентации, GPS и барометра, причем каждый датчик фиксирует параметры с собственной частотой и различной точностью. Это требует обязательной временной синхронизации, позволяющей выровнять записи и сформировать согласованные ряды наблюдений. Далее осуществляется предобработка: удаление шумов, коррекция выбросов и восстановление пропущенных значений с использованием фильтров Калмана, медианных и адаптивных сглаживающих алгоритмов.

Процесс сбора данных с пространственно-временных датчиков смартфона и их последующая обработка представляет собой многоэтапную процедуру, направленную на получение точных, согласованных и информативных параметров, необходимых для построения цифрового паттерна поведения пользователя. На этапе сбора данные поступают от акселерометра, гироскопа, магнитометра, датчиков ориентации, GPS и барометра. Каждый из этих датчиков регистрирует изменения положений, ускорений, высоты и координат в реальном времени. Сбор выполняется либо непрерывно, либо с заданным интервалом, при этом каждая запись снабжается временной меткой. Это позволяет объединять данные разных датчиков в единое временное пространство и отслеживать динамику движений и перемещений пользователя.

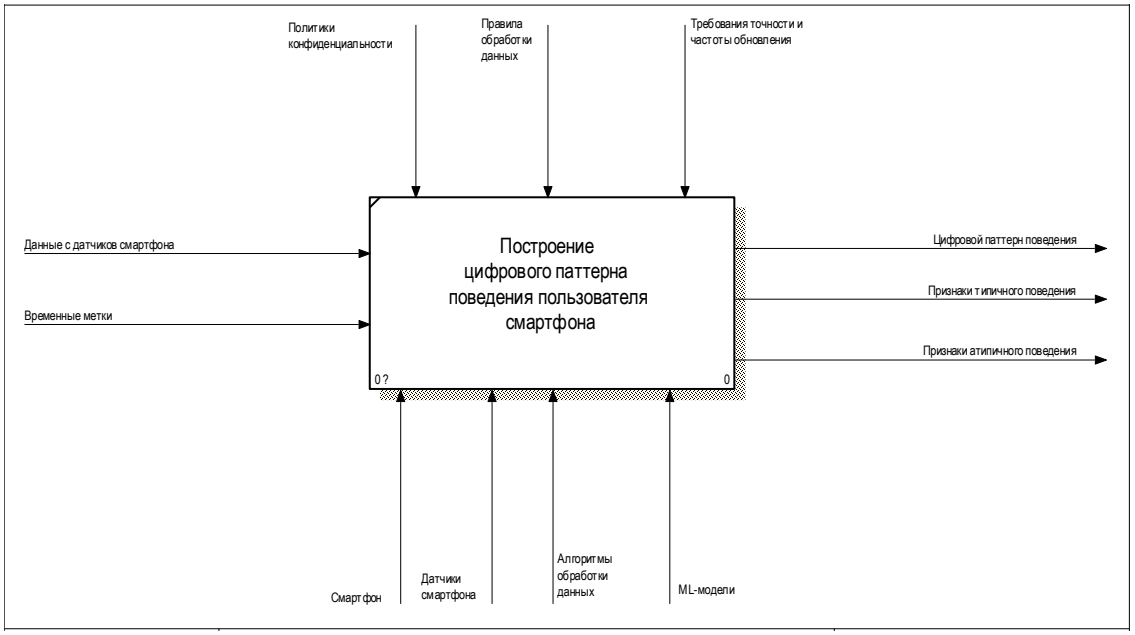


Рисунок 2 – Методика обработки информации с пространственно-временных датчиков смартфона для построения цифрового паттерна поведения пользователя (контекстная диаграмма)

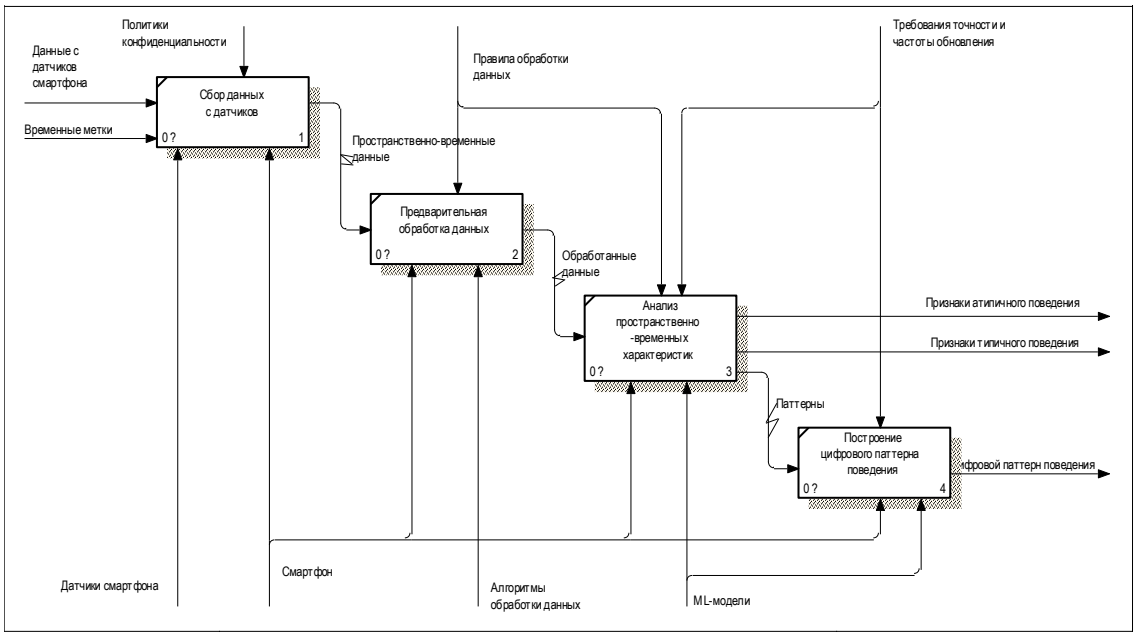


Рисунок 3 – Методика обработки информации с пространственно-временных датчиков смартфона для построения цифрового паттерна поведения пользователя (декомпозиция контекстной диаграммы)

После обработки данных выполняется извлечение признаков, включающее расчет статистических метрик, спектральный анализ для выявления скрытых циклов, определение параметров траектории и характеристик моторики пользователя. На основе этих признаков формируется многомерное поведенческое пространство, в котором методы машинного обучения позволяют выделить устойчивые паттерны, отличающие конкретного пользователя от других. Проводится оценка стабильности паттерна, его воспроизводимости при различных сценариях использования смартфона и устойчивости к внешним помехам.

Результатом методики становится цифровой паттерн поведения – структурированная модель, позволяющая решать задачи идентификации, персонализации и мониторинга активности, обеспечивая высокую точность и надежность анализа пользовательского поведения.

ЗАКЛЮЧЕНИЕ

В результате проведенного исследования была разработана методика обработки информации с пространственно-временных датчиков смартфона, позволяющая формировать цифровой паттерн поведения пользователя. Предложенный подход объединяет сбор, предварительную обработку, анализ и моделирование данных, получаемых с датчиков смартфона, с целью выявления как типичных, так и атипичных паттернов поведения. Использование такой методики обеспечивает возможность детального понимания повседневной активности пользователя, прогнозирования его перемещений и предпочтений, а также выявления отклонений от привычного поведения. Разработка цифровых паттернов поведения на основе пространственно-временных данных открывает перспективы для применения в различных областях, включая персонализированные сервисы, адаптивные интерфейсы и системы безопасности.

Список источников

1. Sela, A. Smartphone use behavior and quality of life: What is the role of awareness? / A. Sela, N. Rozenboim, H. C. Ben-Gal // PLoS ONE. – 2022. – Vol. 17 (3). – Article e0260637. https://doi.org/10.1371/journal.pone.0260637.

2. Wang T. Smartphone Usage Patterns and Sleep Behavior in Demographic Groups: Retrospective Observational Study / T. Wang, A. Seiger, A. Markowetz, I. Andone, K. Błaszkiewicz, T. Penzel // J. Med. Internet Res. – 2025. – Vol. 27. – Article e60423.

3. Павлов, А. В. Анализ данных акселерометра на платформе Tizen / А. В. Павлов, А. Р. Солодовникова, Е. А. Ильюшин, Д. Е. Намиот // International Journal of Open Information Technologies. – 2018. – № 2. – URL: https://cyberleninka.ru/article/n/analiz-dannyh-akselerometra-na-platforme-tizen (дата обращения: 14.05.2025).

4. Тузов, А. Датчики для измерения параметров движения на основе MEM8технологии. Часть 1. Инерциальные датчики средней точности / А. Тузов // Электроника: наука, технология, бизнес. – 2011. – № 1. – С. 46–52.

5. Получение информации с сенсоров устройства Android. – URL: https://android-tools.ru/help/poluchenie-informacii-s-sensorov-ustrojstva-android/ (дата обращения: 14.05.2025).

6. Billieux, J. Can disordered mobile phone use be considered a behavioral addiction? An update on current evidence and a comprehensive model for future research / J. Billieux, P. Maurage, O. Lopez-Fernandez, D. J. Kuss, M. D. Griffiths // Current Addiction Reports. – 2015. – Vol. 2 (2). – P. 156–162.

7. Konok, V. Humans' attachment to their mobile phones and its relationship with interpersonal attachment style / V. Konok, D. Gigler, B. M. Bereczky, Á. Miklósi // Computers in Human Behavior. – 2016. – Vol. 61. – P. 537–47.

8. Ellis, D. A. Digital Traces of behavior within addiction: Response to Griffiths (2017) / D. A.Ellis, L. K. Kaye, T. D. Wilcockson, F. C. Ryding // International Journal of Mental Health and Addiction. – 2018. – Vol. 16 (1). – P. 240–245.

References

1. Sela, A., Rozenboim, N., Ben-Gal, H. C. Smartphone use behavior and quality of life: What is the role of awareness? *PLoS ONE*, 2022, vol. 17 (3), article e0260637. https://doi.org/10.1371/journal.pone.0260637.

2. Wang, T., Seiger, A., Markowetz, A., And one, I., Błaszkiewicz, K., Penzel, T. Smartphone Usage Patterns and Sleep Behavior in Demographic Groups: Retrospective Observational Study. *J. Med. Internet Res.*, 2025, vol. 27, article e60423.

3. Pavlov, A. V., Solodovnikova, A. R., Ilyushin, E. A., Namiot, D. E. Accelerometer data analysis on the Tizen platform. *International Journal of Open Information Technologies*, 2018, no. 2. Available at: https://cyberleninka.ru /article/n/analiz-dannyh-akselerometra-na-platforme-tizen (accessed 14.05.2025) (In Russ.).

4. Tuzov, A. Sensors for measuring motion parameters based on MEM8technology. Part 1. Inertial sensors of medium accuracy. *Electronics:Science, Technology, Business*, 2011, no. 1, pp. 46–52 (In Russ.).

5. *Receiving information from Android device sensors*. Available at: https://android-tools.ru/help/poluchenie-informacii-s-sensorov-ustrojstva-android/ (accessed 14.05.2025) (In Russ.).

6. Billieux, J., Maurage, P., Lopez-Fernandez, O., Kuss, D. J., Griffiths M. D. Can disordered mobile phone use be considered a behavioral addiction? An update on current evidence and a comprehensive model for future research. *Current Addiction Reports*, 2015, vol. 2 (2), pp. 156–162.

7. Konok, V., Gigler, D., Bereczky, B. M., Miklósi, Á. Humans' attachment to their mobile phones and its relationship with interpersonal attachment style. *Computers in Human Behavior*, 2016, vol. 61, pp. 537–547.

8. Ellis, D. A., Kaye, L. K., Wilcockson, T. D., Ryding, F. C. Digital Traces of behavior within addiction: Response to Griffiths (2017). *International Journal of Mental Health and Addiction*, 2018, vol. 16 (1), pp. 240–245.

Статья поступила в редакцию 03.07.2025; одобрена после рецензирования 08.08.2025; принята к публикации 25.10.2025.

The article was submitted 03.07.2025; approved after reviewing 08.08.2025; accepted for publication 25.10.2025.

УДК 004.056+614.2+342.9

МОДЕЛЬ КИБЕРРИСКОВ ЦИФРОВОГО ЗДРАВООХРАНЕНИЯ И РЕГУЛЯТОРНЫЕ ОГРАНИЧЕНИЯ

Ржевская Наталья Витальевна, Северо-Кавказский Федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

аспирант кафедры вычислительной математики и кибернетики, ORCID: 0009-0002-1285-4196, e-mail: natalia070901@gmail.com

Ледян Денис Игоревич, Московский финансово-юридический университет, 115191, Российская Федерация, г. Москва, ул. Серпуховский вал, 17, корп. 1,

аспирант, ORCID: 0009-0005-1462-0635, e-mail: denisledyan@gmail.com

Лапина Мария Анатольевна, Северо-Кавказский Федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

кандидат физико-математических наук, доцент, доцент кафедры вычислительной математики и кибернетики факультета математики и компьютерных наук имени профессора Н.И. Червякова, ORCID: 0000-0001-8117-9142, e-mail: mlapina@ncfu.ru

Лапин Виталий Геннадьевич, Северо-Кавказский Федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

кандидат физико-математических наук, доцент, доцент кафедры вычислительной математики и кибернетики факультета математики и компьютерных наук имени профессора Н.И. Червякова, ORCID: 0000-0002-0611-7002, e-mail: vitlx@yandex.ru

В статье рассматриваются ключевые аспекты электронного здравоохранения с точки зрения доверия пользователей, киберугроз и их влияния на защиту данных. Проведен анализ восприятия электронного здравоохранения населением на основе данных анонимного опроса, в котором участвовали граждане разных возрастных категорий. В ходе анализа выявлены основные опасения и ожидания в области конфиденциальности и безопасности. Представлены точки зрения злоумышленников и управляющих органов на киберугрозы, а также оценена роль законодательства и нормативных актов в обеспечении защиты данных. Особое внимание уделено балансу между развитием технологий и необходимостью их безопасного использования. Работа направлена на понимание потребностей общества в области защиты данных в электронном здравоохранении и предложении возможных решений. В качестве методов исследования применены опрос, контент-анализ существующих нормативных актов и обзор реальных случаев утечек данных. Новизна исследования заключается в комплексном подходе, объединяющем мнение пользователей, анализ угроз и роль регулирующих органов в контексте цифровой медицины. Особое внимание уделено балансу между развитием технологий и необходимостью их безопасного использования.

Ключевые слова: электронное здравоохранение, киберугрозы, защита данных, доверие пользователей, конфиденциальность, информационная безопасность, законодательство, цифровая медицина, кража данных, нормативные акты

DIGITAL HEALTHCARE CYBERRISK MODEL AND REGULATORY RESTRICTIONS

Rzhevskaya Natal'ya V., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

graduate student of the Department of Computational Mathematics and Cybernetics, ORCID: 0009-0002-1285-4196, e-mail: natalia070901@gmail.com

Ledyan Denis I., Moscow University of Finance and Law, 17, bldg. 1 Serpukhovsky Val St., Moscow, 115191, Russian Federation,

postgraduate student, ORCID: 0009-0005-1462-0635, e-mail: denisledyan@gmail.com

Lapina Maria A., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

Cand. Sci. (Physics & Mathematics), Associate Professor, ORCID: 0000-0001-8117-9142, e-mail: mlapina@ncfu.ru

Lapin Vitaly G., North Caucasus Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

Cand. Sci. (Physics & Mathematics), Associate Professor, ORCID: 0000-0002-0611-7002, e-mail: vitlx@yandex.ru

The article examines key aspects of electronic healthcare (e-health) from the perspectives of user trust, cyber threats, and their impact on data protection. An analysis of public perception of e-health was conducted based on anonymous survey data involving citizens from various age groups. The study identified major concerns and expectations related to confidentiality and security. The viewpoints of cybercriminals and regulatory bodies on cyber threats are presented, and the role of legislation and regulatory frameworks in ensuring data protection is assessed. Special attention is given to balancing technological advancement with the necessity for secure implementation. The study aims to understand societal needs regarding data protection in e-health and propose potential solutions. Research methods include a survey, content analysis of existing regulations, and a review of real cases of data breaches. The novelty of this study lies in its comprehensive approach, integrating user opinions, threat analysis, and the role of regulatory bodies in the context of digital healthcare. Particular emphasis is placed on balancing technological progress with the need for its secure use.

Keywords: electronic healthcare, cyber threats, data protection, user trust, confidentiality, information security, legislation, digital medicine, data breaches, regulatory frameworks

ВВЕДЕНИЕ

Электронное здравоохранение (далее – ЭЗ, e-health) представляет собой современный подход к организации медицинских услуг, основанный на применении цифровых технологий для управления данными, улучшения взаимодействия между пациентами и медицинскими учреждениями, а также повышения качества и доступности медицинской помощи. В условиях стремительной цифровизации медицины ЭЗ становится неотъемлемой частью национальных и международных стратегий в области здравоохранения, открывая возможности для инноваций и повышения эффективности медицинских процессов.

Однако активное использование цифровых технологий в медицине сопровождается ростом рисков, связанных с защитой данных. Персональные медицинские данные являются одной из самых ценных категорий информации, что делает их привлекательной целью для злоумышленников. Утечка таких данных может привести к серьезным последствиям, включая нарушение конфиденциальности, мошенничество и дискредитацию медицинских учреждений. При этом доверие пользователей к системам ЭЗ напрямую зависит от уровня их безопасности, что делает вопрос защиты данных в данной области особо актуальной проблемой для исследователей, разработчиков и регуляторов.

Целью данного исследования является изучение отношения населения к электронному здравоохранению, анализ их доверия и опасений в контексте конфиденциальности и безопасности данных. Также исследуются точки зрения злоумышленников и управляющих органов, их роль в формировании угроз и обеспечении защиты данных.

В рамках исследования планируется: проанализировать общественное мнение на основе данных опроса, изучить факторы, определяющие уровень доверия к ЭЗ, проанализировать существующие киберугрозы и их воздействие на безопасность данных, а также оценить роль законодательства и нормативных актов в информационной безопасности.

Таким образом, работа направлена на создание целостного представления о киберугрозах и доверии в ЭЗ, а также разработку рекомендаций для повышения уровня безопасности и доверия к этим системам.

МЕТОДЫ ИССЛЕДОВАНИЯ

В рамках исследования был проведен анонимный опрос среди пользователей систем ЭЗ в период с 14 по 21 ноября 2024 г. [1]. Опрос включал вопросы, связанные с частотой использования цифровых медицинских услуг, уровнем доверия граждан к таким системам и их опасениями относительно защиты персональных и медицинских данных. Респонденты представляли различные возрастные группы, что позволило учесть широкий спектр мнений и определить основные факторы, влияющие на доверие к электронному здравоохранению. География опроса составила преимущественно Ставропольский край, а также Краснодарский край, Республику Дагестан, Астраханскую и Московскую области, Ямало-Ненецкий автономный округ.

Дополнительно был проведен анализ статистических данных о случаях утечек медицинской информации. Исследованы частота подобных инцидентов, их последствия и востребованность медицинских данных у злоумышленников. Особое внимание уделено целям, с которыми данные похищаются, включая их использование в финансовых мошенничествах, шантаже или продаже на черном рынке.

Также были изучены нормативные документы и требования регуляторов, чтобы полноценно понимать принципы защиты информации в электронном здравоохранении. В частности, были проанализированы меры, которые правоохранительные органы применяют для предотвращения утечек и комплексной защиты данных. Для формирования полноценной картины был рассмотрен международный опыт в данной области.

Данный подход исследования был выбран для формирования целостного представления мер защиты данных в области электронного здравоохранения, а также анализа эффективности данных мер с точки зрения доверия пользователей к технологии.

АНАЛИЗ УРОВНЯ ДОВЕРИЯ И ОТНОШЕНИЯ НАСЕЛЕНИЯ К ЭЛЕКТРОННОМУ ЗДРАВООХРАНЕНИЮ

Электронное здравоохранение постепенно становится важной частью современной медицинской инфраструктуры, однако его внедрение сопровождается неоднозначным отношением со стороны населения.

Большинство респондентов выразили положительный ответ о знании понятия электронного здравоохранения, но при этом значительная доля опрошенных имеют только косвенное представление или не имеют понятия о данных технологиях (рис. 1). При этом, как ни удивительно, самыми осведомленными в данном вопросе стали опрошенные в возрасте от 18 до 29 лет.

Рисунок 1 – Уровень осведомленности в понятии ЭЗ

Анализ частоты использования услуг ЭЗ показывает, что большинство опрошенных иногда пользуются или планируют попробовать данные услуги (рис. 2а). Также опрошенные активно используют приложение для отслеживания здоровья (77 % опрошенных) (рис. 2б) и различные электронные устройства для отслеживания показателей здоровья (рис. 2в).

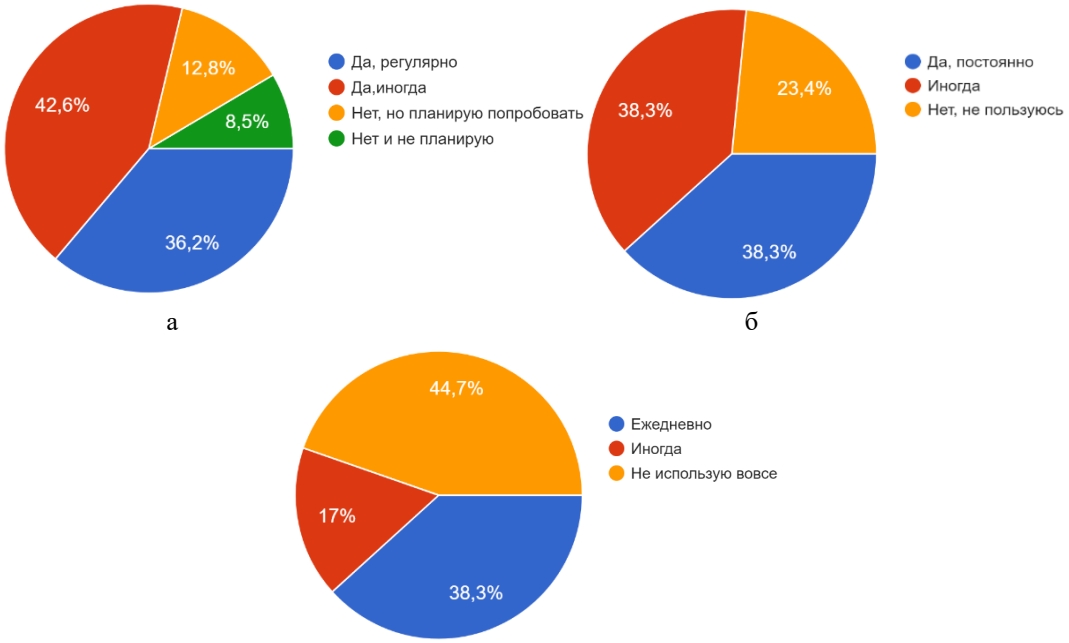


Рисунок 2 – Уровень осведомленности в понятии ЭЗ

Самими востребованными услугами ЭЗ стали онлайн-записи и интернет-консультации (рис. 3).

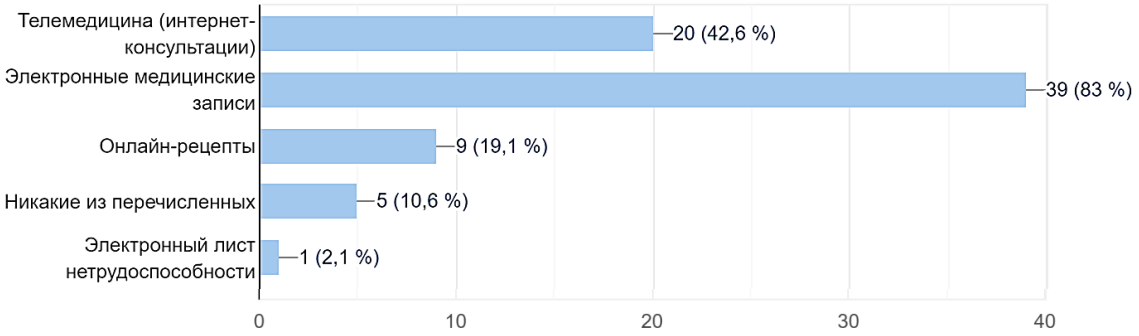


Рисунок 3 – Востребованность услуг электронного здравоохранения

Анализ общественного мнения показывает, что многие граждане положительно воспринимают возможности ЭЗ (рис. 4). Так, 85 % респондентов оценили свое отношение к ЭЗ как положительное. Средняя оценка отношения к электронному здравоохранению составила 4,49.

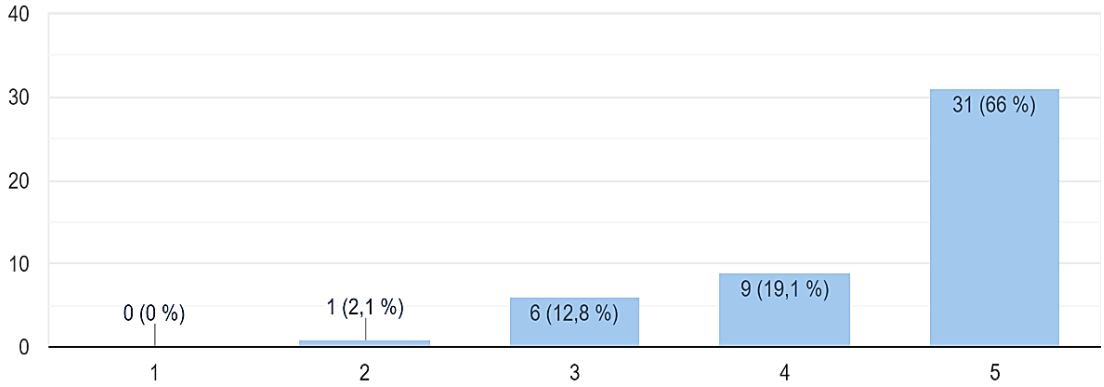


Рисунок 4 – Уровень отношения к электронному здравоохранению (где 5 – очень положительно, 1 – очень отрицательно)

К основным плюсам данных систем респонденты отнесли удобный доступ к медицинским услугам, дистанционные консультации и оперативное хранение данных (рис. 5). Удобство и скорость доступа к информации были выделены около 75 % участников как главные причины использования ЭЗ.

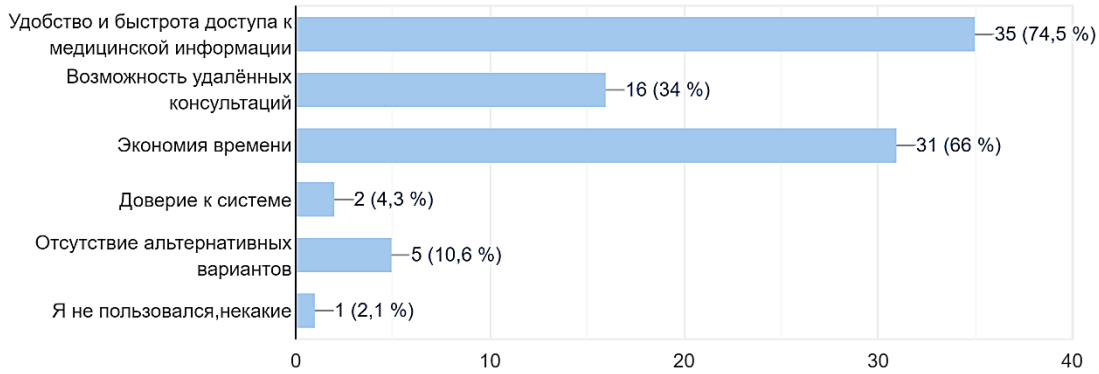


Рисунок 5 – График основных причин использования ЭЗ

Тем не менее, наряду с преимуществами, значительная часть населения выражает опасения по поводу безопасности личной информации и недостатка контроля над процессами обработки данных (рис. 7). Помимо преимуществ, выявлены ключевые барьеры, снижающие доверие к ЭЗ. Например, 62 % участников опроса проявляют осторожность в доверии врачам, консультирующим онлайн (рис. 6), и около 9 % полностью избегают таких услуг.

Рисунок 6 – Основные опасения по отношению к ЭЗ

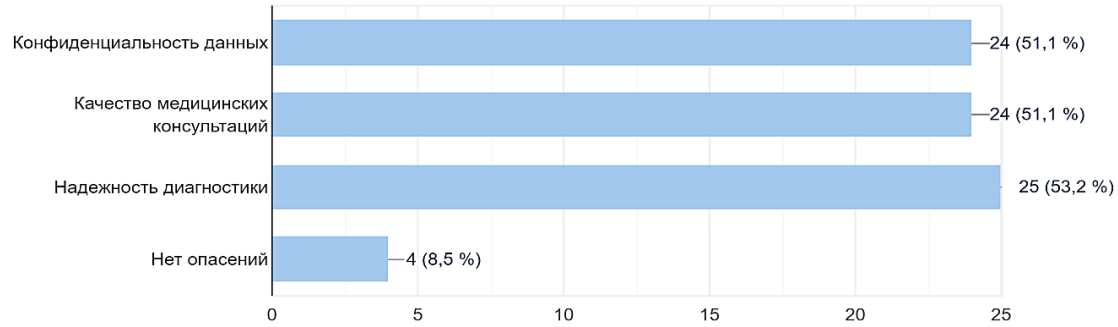


Рисунок 7 – Основные опасения по отношению к ЭЗ

Стоит отметить, что большая часть опрошенных отметили, что в случае утечки персональных медицинских данных, уровень доверия к такой технологии исчезнет полностью или снизиться существенно (рис. 8).

Рисунок 8 – Отношение к ЭЗ в случае утечки персональных медицинских данных

При этом основные опасения связаны с раскрытием конфиденциальной информации, а также возможности появления у злоумышленника необходимых данных для различного рода шантажа (рис. 9).

Рисунок 9 – Основные опасения по отношению к ЭЗ

Особенно чувствительной оказалась тема шифрования данных, связанных с психиатрическими и другими конфиденциальными записями, что поддержали более 70 % респондентов.

Доверие к системам ЭЗ формируется под влиянием множества факторов. Ключевыми из них являются прозрачность использования данных, наличие эффективных механизмов защиты и информированность граждан о том, как работают системы ЭЗ.

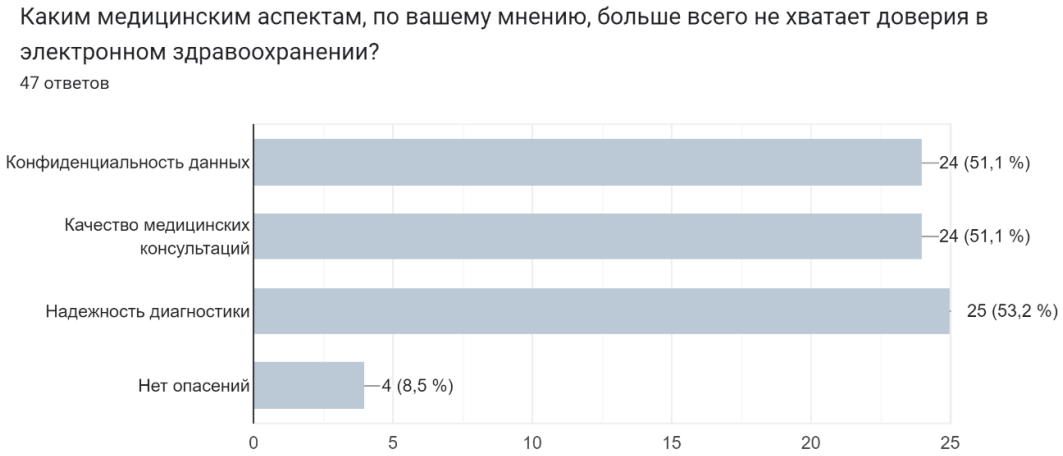


Рисунок 10 – Основные опасения по отношению к ЭЗ

Основные опасения связаны с возможностью утечек данных (65 %) и сложностью интерфейсов систем (35 %), что делает необходимым повышение удобства их использования и прозрачности работы. Барьеры на пути формирования доверия включают частые сообщения о кибератаках, отсутствие уверенности в компетентности разработчиков и регуляторов, а также опасения по поводу несанкционированного доступа к медицинским данным.

Интересным аспектом исследования стало отношение граждан к межорганизационному обмену медицинскими данными. Подавляющее большинство респондентов (66 %) поддерживают эту идею при условии строгого соблюдения мер безопасности (рис. 11). Это подчеркивает важность разработки надежных механизмов передачи данных между учреждениями здравоохранения.

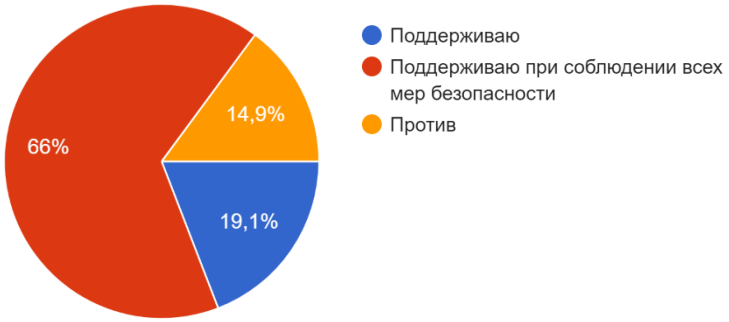


Рисунок 11 – Отношение к межорганизационному обмену медицинскими данными

При оценке существующих регуляторных механизмов защиты данных 28 % участников отметили их удовлетворительный уровень, однако 41 % затруднились с ответом, что указывает на недостаток информированности населения о текущих мерах и нормах безопасности (рис. 12). Способность отметить хорошее знание о методах и мерах защищенности персональных данных в электронном здравоохранении смогли отметить только 30 % опрошенных (рис. 13). Интересным оказался факт, что значительная часть респондентов даже не предполагают, что данная сфера защищается каким-либо образом в нормативно-правовых актах (рис. 14).

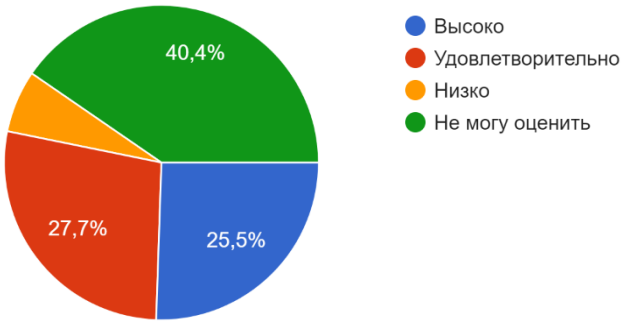


Рисунок 12 – Оценка современного состояния защищенности данных при использовании ЭЗ

Рисунок 13 – Оценка знаний мер и методов защиты данных в ЭЗ

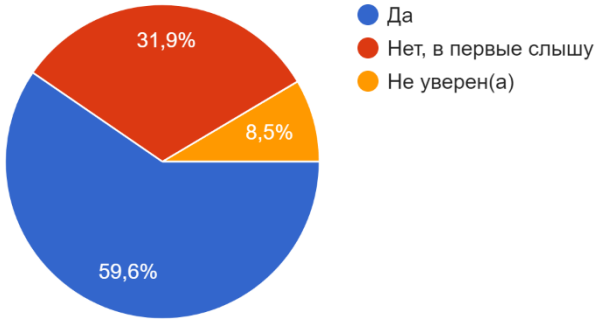


Рисунок 14 – Оценка знаний существования мер защиты ЭЗ

Опрос респондентов о необходимых мерах защиты данных в электронном здравоохранении выявил несколько ключевых аспектов. Большинство участников частично осведомлены о законодательной защите данных, но сохраняется недоверие к существующим механизмам. Почти 60 % респондентов считают электронное здравоохранение более уязвимым, чем традиционные методы, что свидетельствует о необходимости повышения прозрачности и внедрения более надежных систем. При этом пользователи ожидают обязательного информирования о доступах к их данным и считают важным применять биометрические и двухфакторные методы аутентификации.

Также анализ опроса показал, что ряд респондентов считают одним из эффективных методов повышения защиты данных обучение медицинского персонала в области информационной безопасности. Стоит отметить, что 80 % опрошенных выступают за обязательное или желательное обучения персонала при работе с чувствительными данными. Значительная часть респондентов выразили желание, чтобы медицинский персонал, обрабатывающий психиатрические и(или) генетические данные, обучался в области защиты данных в обязательном порядке.

Анализ опроса показал, что пользователи готовы работать с электронным здравоохранением. Однако не все опрошенные считают данную систему прозрачной и понятной. Респонденты готовы довериться данной технологии, только при условии повышения надежности и минимизации рисков утечки чувствительной конфиденциальной информации. Также мерами защиты данных респонденты назвали совершенствование законодательства в данной области, применение передовых средств защиты информации и криптографических средств, а также обучение персонала. Стоит отметить, что значительная часть опрошенных выразили желание повысить осведомленность в области защиты медицинских данных, назвав это одним из ключевых механизмов повышения доверия к данным системам.

РОЛЬ УПРАВЛЯЮЩИХ ОРГАНОВ В ОБЕСПЕЧЕНИИ БЕЗОПАСНОСТИ ЭЗ

Роль управляющих органов в обеспечении безопасности данных ЭЗ является основополагающим элементом развития данной сферы, особенно в условиях стремительной цифровизации медицинских услуг. Электронное здравоохранение как интегративный инструмент современной медицины позволяет расширить доступ к медицинским услугам, однако одновременно с этим увеличиваются риски утечек персональной информации, киберугроз и атак на информационные системы, что делает вопросы регулирования и защиты данных особенно актуальными. В условиях нарастающей значимости цифровых технологий безопасность становится краеугольным камнем для укрепления доверия пользователей.

Управляющие органы стремятся создать условия, при которых развитие технологий будет безопасным, а уровень доверия к цифровым медицинским платформам – высоким. Российское законодательство, формируемое на основе Федерального закона № 152-ФЗ «О персональных данных» и Федерального закона № 323-ФЗ «Об основах охраны здоровья граждан», за последние годы претерпело значительные изменения, направленные на повышение уровня защиты персональной информации. Так, статистические данные показывают, что за пять лет количество нормативных актов, касающихся безопасности данных, увеличилось почти на 37 % [19], что свидетельствует о возрастающем внимании к данному вопросу.

Ключевыми способами повышения защищенности данных являются реализация обязательного шифрования данных, введение строгой аутентификации пользователей и усиление контроля доступа к медицинской информации. Так, значительным шагом в регулировании обработки и хранения медицинских данных стал Приказ № 965 Минздрава России. Данный приказ позволил систематизировать процессы обработки данных в области как медицины в целом, так и электронного здравоохранения в частности.

Стоит отметить, что российское законодательство в сфере здравоохранения продолжает развиваться, делая упор на безопасность персональных данных. Ужесточение наказаний за нарушения в сфере защиты данных стало важным изменением в российском законодательстве. Недавние поправки к статье 13.11 КоАП РФ предусматривают штрафы до 20 миллионов рублей за утечку биометрических данных, а при повторных случаях – до 3 % от годового дохода компании. Однако компании, соблюдающие закон, могут получить снижение штрафов в 10 раз. Это говорит о том, что российская правовая база стремится к созданию более устойчивой системы защиты данных и норм регулирования. Решение этих задач поможет создать надежную защиту информации, уменьшить риск утечек и повысить доверие пациентов.

Международный опыт демонстрирует более высокий уровень стандартизации в области защиты данных. Например, в странах Европейского союза действует Общий регламент о защите данных (GDPR), который определяет строгие требования к обработке персональной информации, включая медицинские данные. Согласно отчетам Европейской комиссии, с момента внедрения GDPR количество утечек данных в здравоохранении сократилось на 12 % [18]. В США действует Закон HIPAA, регулирующий стандарты безопасности в медицинской сфере, что позволило, полагаясь на отчеты HIPAA Journal [17], за последние три года снизить количество инцидентов несоответствия требованиям на 19 %. Все это свидетельствует, что такой подход к защите данным является наиболее эффективным.

Однако стоит отметить, что регулирующие органы нередко сталкиваются с проблемой между конфиденциальностью и доступностью. Так, чрезмерное регулирование, с одной стороны, повышает защищенность персональных данных, а с другой стороны, тормозит внедрения новых технологий в связи со сложностью соответствия с всеми необходимыми требованиями. Отчет [21] за 2023 г. показал, что убыток от кибератак в сфере здравоохранения превысил 9,2 миллионов долларов. Также стоит отметить, что большая часть утечек произошла по вине недостаточной защищенности систем здравоохранения.

Защита медицинской информации требует комплексного подхода, включающего использование современных технологий, направленных на предотвращение утечек, взломов и краж данных. Согласно исследованию HIMSS Analytics, только 25 % руководителей клиник обеспечивают безопасность данных в соответствии с установленными стандартами. Однако в ряде стран уже внедрены эффективные решения.

Например, госпиталь Albert Schweitzer Ziekenhuis в Нидерландах применяет одноразовые токены и облачные серверы для аутентификации, что обеспечивает высокий уровень конфиденциальности. Швеция разработала национальную систему «Электронная медицинская карта», использующую двухфакторную идентификацию через смарт-карты, логины и цифровые подписи, что позволяет контролировать действия сотрудников и минимизировать риски утечек.

В Канаде смарт-карты используются для доступа к медицинским записям, что способствует улучшению качества работы медицинских учреждений и снижению вероятности инцидентов. Современные технологии, такие как системы для смартфонов на основе Bluetooth и адаптеров

для электронных удостоверений, обеспечивают новый уровень безопасности и удобства. Они позволяют медицинским специалистам удаленно подписывать документы и выписывать рецепты. Такие решения активно применяются в развитых странах и продолжают совершенствоваться, способствуя надежной защите медицинской информации.

Таким образом, роль регулирующих органов заключается не только в создании условий для безопасного развития электронного здравоохранения, но и в обеспечении механизмов контроля за исполнением требований. Усиление ответственности за нарушение норм, вкупе с внедрением передовых технологий защиты, способствует формированию надежной системы, минимизирующей киберугрозы и укрепляющей доверие пользователей.

ЗАКЛЮЧЕНИЕ

Электронное здравоохранение остается ключевой частью цифровизации здравоохранения, но, как показывает практика, его восприятие населением неоднозначно. Так, одной из важнейших проблем остается баланс между осознанием преимуществ цифровых медицинских технологий и опасениями относительно их безопасности и приватности. Из этого следует, что для повышения доверия к электронному здравоохранению необходимо проводить активные образовательные кампании, а также усиливать защиту данных. Одновременно важно обеспечить взаимодействие между разработчиками, регулирующими органами и обществом.

Так, исследование показывает, что более молодые группы населения, обладающие большей технологической грамотностью, чаще выражают интерес к использованию электронных медицинских услуг. Однако лица старшего возраста, как правило, воспринимают такие технологии с настороженностью. Это связано как с их опасениями по поводу сложности интерфейсов, так и с боязнью утечки данных. Социальные стереотипы, например убеждение в уязвимости цифровых систем, также усиливают настороженность.

Кроме того, одной из главных причин повышенного интереса злоумышленников к ЭЗ является высокая ценность медицинской информации. Такие данные, как диагнозы, история болезней, результаты анализов, номера страховых полисов, представляют значительный интерес для преступников, так как могут быть использованы для финансовых махинаций или шантажа. Так, утечка информации часто используется для фальсификации медицинских документов, создания ложных страховок или мошенничества с лекарствами.

Учитывая рост числа атак, включая фишинг и внедрение вредоносных программ, можно заключить, что обеспечение защиты медицинских данных должно быть одной из первоочередных задач всех участников системы электронного здравоохранения. Важное значение здесь имеет как совершенствование нормативной базы, так и повышение квалификации медицинского персонала. Одновременно необходимо укреплять международное сотрудничество для регулирования вопросов трансграничной передачи данных.

Таким образом, эффективная защита медицинской информации требует комплексного подхода, который должен включать регулярное обновление технологий, создание единых стандартов безопасности и четкое выполнение всех установленных требований. Подобные меры не только снизят риски утечек и кибератак, но и повысят доверие населения к цифровым медицинским платформам.

Список источников

1. Опрос об отношении к электронному здравоохранению // Google Form. – URL: https://docs.google.com /forms/d/1aIMQrIp-O9AFIHQLJoWggE1a9y3PP2SyDagmOZiGyK4/edit#responses (дата обращения: 24.11.2024).

2. Kaspersky: Кибербезопасность в здравоохранении: где болезнь, где болезнь роста. – URL: https://www.kaspersky.ru/blog/healthcare-safeguarding-data/4474/.

3. Иванова, А. А. Применение big data в сфере здравоохранения: Российский и зарубежный опыт / А. А. Иванова // Научные записки молодых исследователей. – 2020. – № 5. – URL: https://cyberleninka.ru/article/n /primenenie-big-data-v-sfere-zdravoohraneniya-rossiyskiy-i-zarubezhnyy-opyt (дата обращения: 01.11.2023).

4. Андриянова, Е. А. Проблемы формирования системы электронного здравоохранения в России / Е. А. Андриянова, Н. В. Гришечкина // Здравоохранение российской федерации. – 2012. – № 6. – С. 27–30.

5. Информатизация здоровья. Менеджмент защиты информации в здравоохранении по ИСО/МЭК 27002.

6. Концепция информационной безопасности в сфере здравоохранения (утв. протоколом президиума Правительственной комиссии по цифровому развитию, использованию информационных технологий для улучшения качества жизни и условий ведения предпринимательской деятельности от 10.03.2022 № 7).

7. Российская Федерация. Законы. О безопасности критической информационной инфраструктуры Российской Федерации. федер. закон от 26.07.2017 № 187-ФЗ // Справочно-правовая система «Консультант Плюс». – URL: http://www.consultant.ru/document/cons_doc_LAW_220885/.

8. Постановление Правительства РФ от 8 февраля 2018 г. № 127 «Об утверждении Правил категорирования объектов критической информационной инфраструктуры Российской Федерации, а также перечня показателей критериев значимости объектов критической информационной инфраструктуры Российской Федерации и их значений».

9. ГОСТ Р ИСО 27799-2015. Информатизация здоровья. Менеджмент защиты информации в здравоохранении по ИСО/МЭК 27002: утв. приказом Федерального агентства по техническому регулированию и метрологии от 28 декабря 2015 г. № 2219-ст. – Москва, 2016.

10. Российская Федерация. Законы. О персональных данных: Федеральный закон от 27.07.2006 № 152-ФЗ // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document /cons_doc_LAW_61801/.

11. Российская Федерация. Методический документ. Методика оценки угроз безопасности информации: утв. ФСТЭК России 05.02.2021 // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document/cons_doc_LAW_378330/.

12. Российская Федерация. Приказ ФСТЭК России от 18.02.2013 № 21 (ред. от 14.05.2020) «Об утверждении Состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных» : зарегистрировано в Минюсте России 14.05.2013 № 28375 // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document/cons_doc_LAW_146520/.

13. Российская Федерация. Приказ ФСТЭК России от 11.02.2013 № 17 (ред. от 28.05.2019) «Об утверждении Требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах» : зарегистрировано в Минюсте России 31.05.2013 № 28608) (с изм. и доп., вступ. в силу с 01.01.2021) // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document/cons_doc_LAW_147084/.

14. Российская Федерация. Законы. Об информации, информационных технологиях и о защите информации: Федеральный закон от 27.07.2006 № 149-ФЗ // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document/cons_doc_LAW_61798/.

15. Российская Федерация. Законы. Об основах охраны здоровья граждан в Российской Федерации: Федеральный закон от 21.11.2011 № 323-ФЗ // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document/cons_doc_LAW_121895/.

16. Российская Федерация. Постановление Правительства. Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных: Постановление Правительства РФ от 1 ноября 2012 г. № 1119 // Справочно-правовая система «Консультант Плюс». – URL: https://www.consultant.ru/document/cons_doc_LAW_137356/8c86cf6357879e861790a8a7ca8bea4227d56c72/.

17. Healthcare Cybersecurity // The HIPAA Journal. – URL: https://www.hipaajournal.com/category/healthcare-cybersecurity/ (дата обращения: 13.11.2024).

18. Highlights of the Data Protection Commission's 2022 Annual Report // mccannfitzgerald. – URL: https://www.mccannfitzgerald.com/knowledge/data-privacy-and-cyber-risk/highlights-of-the-data-protection-commissions-2022-annual-report (дата обращения: 13.11.2023).

19. Быстро и экстремально: количество нормативных актов в сфере ИБ за 2022 год выросло на четверть // Servernews. – URL: https://servernews.ru/1084525 (дата обращения: 13.11.2024).

20. Нормативные правовые акты в сфере информационной безопасности и цифровой экономики по итогам 2021–2022 года // InfoWatch. – URL: https://www.infowatch.ru/sites/default/files/analytics/files /zakonodatelstvo-v-sfere-informatsionnoy-bezopasnosti-i-tsifrovoy-ekonomiki.pdf (дата обращения: 16.11.2023).

21. Healthcare Industry To Spend $125 Billion On Cybersecurity From 2020 to 2025 // cybersecurityventures. – URL: https://cybersecurityventures.com/healthcare-industry-to-spend-125-billion-on-cybersecurity-from-2020-to-2025/ (дата обращения: 16.11.2024).

22. 5 наиболее популярных угроз для индустрии здравоохранения // SecurityLab. – URL: https://www.securitylab.ru/analytics/497623.php (дата обращения: 25.11.2024).

23. Цифровой диагноз: как кибератака парализовала здравоохранение США // SecurityLab. – URL: https://www.securitylab.ru/news/551250.php (дата обращения: 25.11.2024).

24. Лаборатория Касперского представила главные мишени хакеров в 2024 году // SecurityLab. – URL: https://www.securitylab.ru/news/548040.php (дата обращения: 25.11.2024).

25. Эксперты Positive Technologies назвали главные угрозы веб-приложений и оценили рынок WAF в России // SecurityLab. – URL: https://www.securitylab.ru/news/507953.php (дата обращения: 25.11.2024).

26. Эпидемия RAT: исследователи отмечают стремительный рост активности троянов удалённого доступа // SecurityLab. – URL: https://www.securitylab.ru/news/543268.php (дата обращения: 25.11.2024).

27. Утечки конфиденциальных данных и документов из сферы здравоохранения // SecurityLab. – URL: https://www.securitylab.ru/blog/company/EveryTag/352284.php (дата обращения: 25.11.2024).

28. Случаи утечки информации в медицине // SearchInform. – URL: https://searchinform.ru/analitika-v-oblasti-ib/utechki-informatsii/sluchai-utechki-informacii/sluchai-utechki-informacii-v-medicine/ (дата обращения: 25.11.2024).

References

1. Survey on attitudes towards e-healthcare. *Google Form.* Available at: https://docs.google.com/forms /d/1aIMQrIp-O9AFIHQLJoWggE1a9y3PP2SyDagmOZiGyK4/edit#responses (accessed 24.11.2024) (In Russ.).

2. Kaspersky: Cybersecurity in Healthcare: Where the Disease Is, Where the Disease Grows. Available at: https://www.kaspersky.ru/blog/healthcare-safeguarding-data/4474/ (In Russ.).

3. Ivanova, A. A. Application of Big Data in Healthcare: Russian and Foreign Experience. *Scientific Notes of Young Researchers*, 2020, no. 5. Available at: https://cyberleninka.ru/article/n/primenenie-big-data-v-sfere-zdravoohraneniya-rossiyski-i-zarubezhnyy-opyt (accessed 01.11.2023) (In Russ.).

4. Andrianova, E. A., Grishechkina, N. V. Problems of forming the healthcare system in Russia. *Healthcare of the Russian Federation*, 2012, no. 6, pp. 27–30 (In Russ.).

5. *Informatization of health. Information security management in healthcare according to ISO/IEC 27002* (In Russ.).

6. *Concept of information security in healthcare (approved by the minutes of the Presidium of the Government Commission on Digital Development, the use of information technologies to improve the quality of life and business conditions dated 10.03.2022 No. 7)* (In Russ.).

7. Russian Federation. Laws. On the Security of Critical Information Involvement of the Russian Federation. Law of July 26, 2017 No. 187-FZ. *Reference and Legal System "Consultant Plus".* Available at: http://www.consultant.ru/document/cons_doc_LAW_220885/ (In Russ.).

8. *Resolution of the Russian Federation Economy of February 8, 2018 No. 127 "On Approval of the Rules for Categorizing Objects of the Critical Information Industry of the Russian Federation, as well as Clarifying the Innovativeness Indicators of Objects of the Critical Information Industry of the Russian Federation and Their Results"* (In Russ.).

9. *GOST R ISO 27799-2015. Health informatization. Information security management in healthcare according to ISO/IEC 27002: approved. by Order of the Federal Agency for Technical Regulation and Metrology dated December 28, 2015 No. 2219-st.* Moscow, 2016 (In Russ.).

10. Russian Federation. Laws. On Political Data: Federal Law of July 27, 2006, No. 152-FZ. *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_61801/ (In Russ.).

11. Russian Federation. Methodological Document. Information Security Threat Assessment Methodology: approved by the FSTEC of Russia on February 5, 2021. *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_378330/ (In Russ.).

12. Russian Federation. Order of the Federal Service for Technical and Export Control of Russia dated February 18, 2013, No. 21 (as amended on May 14, 2020) "On Approval of the Composition and Content of Organizational and Technical Measures to Ensure the Security of Financial Data During Their Processing in the Ministry of Justice of Russia" dated May 14, 2013, no. 28375. *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_146520/ (In Russ.).

13. Russian Federation. Order of the Federal Service for Technical and Export Control of Russia dated February 11, 2013, No. 17 (as amended on May 28, 2019). On approval of Requirements for the protection of information that do not guarantee a state secret, included in the state information system: Registered in the Ministry of Justice of Russia on 31.05.2013 N 28608) (as amended and supplemented, entered into force on 01.01.2021). *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_147084/ (In Russ.).

14. Russian Federation. Laws. On information, information technology and on the protection of information: Federal Law of 27.07.2006 No. 149-FZ. *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_61798/ (In Russ.).

15. Russian Federation. Laws. On the Fundamentals of Health Protection of Citizens in the Russian Federation: Federal Law of November 21, 2011 No. 323-FZ. *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_121895/ (In Russ.).

16. Russian Federation. Resolution on technologies. On approving requirements for the protection of electronic data when processing them for medical information purposes: Resolution of the Government of the Russian Federation of November 1, 2012 No. 1119. *Reference and Legal System "Consultant Plus".* Available at: https://www.consultant.ru/document/cons_doc_LAW_137356/8c86cf6357879e861790a8a7ca8bea4227d56c72/ (In Russ.).

17. Cybersecurity in Healthcare. *HIPAA Journal.* Available at: https://www.hipaajournal.com/category/healthcare-cybersecurity/ (accessed 11/13/2024).

18. Highlights of the Data Protection Commission's 2022 Annual Report. *mccannfitzgerald.* Available at: https://www.mccannfitzgerald.com/knowledge/data-privacy-and-cyber-risk/highlights-of-the-data-protection-commissions-2022-annual-report (accessed 13.11.2023).

19. Fast and Extreme: The Number of Regulatory Acts in the Sphere of Information Security Increased by a Quarter in 2022. *Servernews.* Available at: https://servernews.ru/1084525 (accessed 13.11.2024) (In Russ.).

20. Regulatory Legal Acts in the Sphere of Information Security and the Digital Economy Based on the Results of 2021–2022. *InfoWatch.* Available at: https://www.infowatch.ru/sites/default/files/analytics/files/zakonodatelstvo-v-sfere-informatsionnoy-bezopasnosti-i-tsifrovoy-ekonomiki.pdf (accessed 16.11.2023) (In Russ.).

21. Healthcare Industry to Spend $125 Billion on Cybersecurity from 2020 to 2025. *cybersecurityventures.* Available at: https://cybersecurityventures.com/healthcare-industry-to-spend-125-billion-on-cybersecurity-from-2020-to-2025/ (accessed 16.11.2024).

22. The 5 Most Popular Threats to the Healthcare Industry. *SecurityLab.* Available at: https://www.securitylab.ru/analytics/497623.php (accessed 25.11.2024) (In Russ.).

23. Digital Diagnosis: How a Cyberattack Paralyzed US Healthcare. *SecurityLab.* Available at: https://www.securitylab.ru/news/551250.php (accessed 25.11.2024) (In Russ.).

24. Kaspersky Lab Presents Hackers' Main Targets in 2024. *SecurityLab.* Available at: https://www.securitylab.ru/news/548040.php (accessed 25.11.2024) (In Russ.).

25. Positive Technologies Experts Name the Main Web Application Threats and Assess the WAF Market in Russia. *SecurityLab.* Available at: https://www.securitylab.ru/news/507953.php (accessed 25.11.2024) (In Russ.).

26. RAT Epidemic: Researchers Note Rapid Increase in Remote Access Trojan Activity. *SecurityLab.* Available at: https://www.securitylab.ru/news/543268.php (accessed 25.11.2024) (In Russ.).

27. Leaks of Confidential Data and Documents from the Healthcare Sector. *SecurityLab.* Available at: https://www.securitylab.ru/blog/company/EveryTag/352284.php (accessed 25.11.2024) (In Russ.).

28. *Cases of Information Leaks in Medicine. SearchInform.* Available at: https://searchinform.ru/analitika-v-oblasti-ib/utechki-informatsii/sluchai-utechki-informacii/sluchai-utechki-informacii-v-medicine/ (accessed 25.11.2024) (In Russ.).

Статья поступила в редакцию 21.11.2025; одобрена после рецензирования 09.12.2025; принята к публикации 09.12.2025.

The article was submitted 21.11.2025; approved after reviewing 09.12.2025; accepted for publication 09.12.2025.

УДК 62:389: 398.14: 53.08

ОЦЕНКА МЕТОДОМ МОНТЕ-КАРЛО НЕОПРЕДЕЛЕННОСТИ ПРИ ПОСТРОЕНИИ АППРОКСИМИРУЮЩЕЙ КРИВОЙ ПО ЭКСПЕРИМЕНТАЛЬНЫМ ДАННЫМ

Шаронов Павел Александрович, Саратовский государственный технический университет, 410054, Российская Федерация, г. Саратов, ул. Политехническая, 77,
аспирант, ORCID: 0009-0003-1568-242X, e-mail: stalker-scharonov@mail.ru
Львов Алексей Арленович, Саратовский государственный технический университет, 410054, Российская Федерация, г. Саратов, ул. Политехническая, 77,
доктор технических наук, профессор, ORCID: 0000-0003-4270-7867, e-mail: alvova@mail.ru

Представлена процедура Монте-Карло для вычисления совместного распределения вероятности, которое следует приписывать коэффициентам кривой, аппроксимирующей набор экспериментальных точек некоторой функциональной зависимости. Аппроксимация кривой проводится на основе имеющихся экспериментальных данных, а также иной доступной исследователю информации. Предлагаемая процедура полностью согласуется с подходом, изложенным в международных и отечественных стандартах. Она заключается в оценке совместного распределения вероятности входных величин путем построения математической модели процесса измерения, с помощью которой определяются коэффициенты. Обычно модель получается методом наименьших квадратов, который в данном случае модифицируется с учетом специфики решаемой задачи оценивания. Однако возможно применение и других критериев аппроксимации. Примеры иллюстрируют использование данной процедуры.

Ключевые слова: аппроксимация, неопределенность измерения, метод Монте-Карло, проверка согласованности, корреляция, распределение, программное отклонение, распределение, распределение вероятности

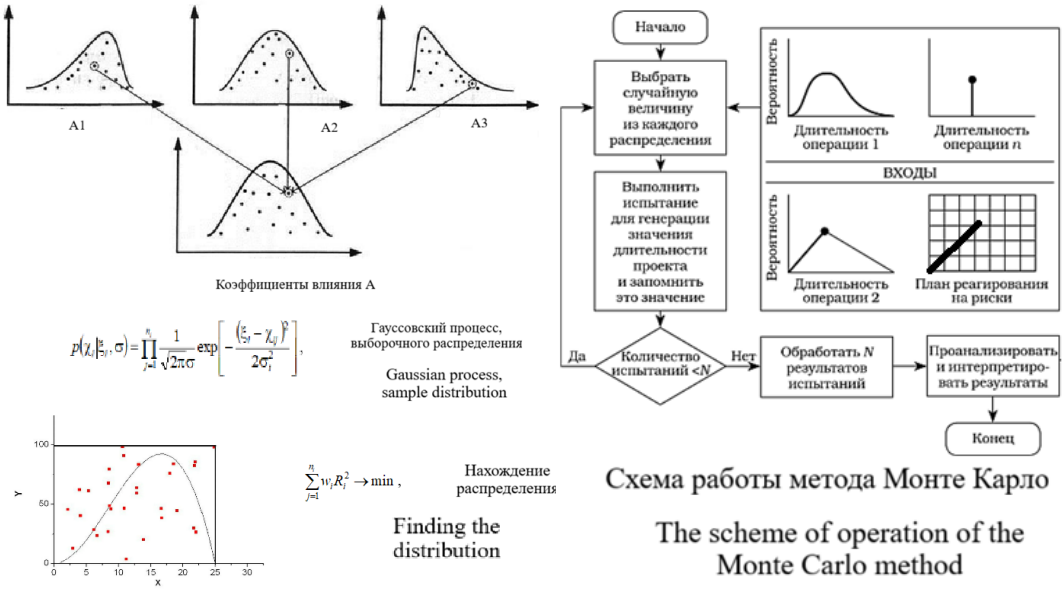
MONTE CARLO ESTIMATION OF UNCERTAINTY WHEN CONSTRUCTING AN APPROXIMATING CURVE FROM EXPERIMENTAL DATA

Sharonov Pavel A., Saratov State Technical University, 77 Politekhnicheskaya St., Saratov, 410054, Russian Federation,
postgraduate student, ORCID: 0009-0003-1568-242X., e-mail: stalker-scharonov@mail.ru
Lvov Alexey A., Saratov State Technical University, 77 Politekhnicheskaya St., Saratov, 410054, Russian Federation,
Doct. Sci. (Engineering), Professor, ORCID: 0000-0003-4270-7867, e-mail: alvova@mail.ru

A Monte Carlo procedure is presented for calculating the joint probability distribution that should be assigned to the coefficients of a curve that approximates a set of experimental points of a certain functional dependence. The curve approximation is based on the available experimental data and other information available to the researcher. The proposed procedure is fully consistent with the approach outlined in international and national standards. It involves estimating the joint probability distribution of the input variables by constructing a mathematical model of the measurement process, which is used to determine the coefficients. Usually, the model is obtained by the least squares method, which in this case is modified to take into account the specifics of the estimation problem. However, other approximation criteria can also be used. The examples illustrate the use of this procedure.

Keywords: approximation, measurement uncertainty, Monte Carlo method, consistency check, correlation, distribution, programmatic deviation, distribution, probability distribution

Graphical annotation (Графическая аннотация)



ВВЕДЕНИЕ

Задача аппроксимации кривой по экспериментальным данным является типичной для многих технических и научных приложений. Она состоит в поиске вектора оценок $\mathbf{a} = (a_1,\ldots,a_n)^T$ неизвестных коэффициентов $\mathbf{A} = (A_1,\ldots,A_n)^T$ (^{T} – оператор транспонирования матрицы) некоторой функции одной переменной $y=F(x, \mathbf{a})$, чтобы получаемая кривая как можно лучше аппроксимировала набор экспериментальных точек $\{x_1, y_1\}$, …, $\{x_m, y_m\}$, при этом число точек m превышает число неизвестных коэффициентов n ($m > n$). Эти точки считаются оценками неизвестных истинных значений $\mathbf{X}=(X_1,\ldots,X_m)^T$ и $\mathbf{Y}=(Y_1,\ldots,Y_m)^T$. Независимо от их реального физического смысла, компоненты этих экспериментальных векторов можно рассматривать как входные случайные переменные в декартовой системе координат, имеющих в общем виде совместное распределение вероятностей $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$. Это распределение должно быть определено или постулировано для данных векторов на основе всей доступной априорной информации о них.

Аналогично коэффициенты **A** тоже являются случайными величинами, подчиняющимися закону распределения, задаваемому совместным распределением вероятности $f_{\mathbf{A}}(\boldsymbol{\alpha})$, которое получается из распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ и из математической модели измерения [1], которая связывает входные координаты с коэффициентами.

Эта модель выводится из критерия, по которому проводится аппроксимация, определяющего, в каком смысле аппроксимирующая кривая должна проводиться «как можно ближе» к имеющимся наборам данных, и позволяющего выразить коэффициенты **A** в неявном виде с помощью некоторой системы уравнений $G(\mathbf{X}, \mathbf{Y}, \mathbf{A}) = 0$. Распределение $f_{\mathbf{A}}(\boldsymbol{\alpha})$ получается из этой системы путем распространения распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$. Значения оценок **a** могут затем интерпретироваться как средние значения маргинальных частных распределений компонентов вектора $f_{A_i}(\alpha_i)$.

Иногда коэффициенты аппроксимирующей кривой сами являются искомыми величинами. В других случаях может потребоваться определение распределения вероятности $f_Z(\boldsymbol{\zeta})$, описывающего случайную измеряемую величину Z, которая каким-либо образом зависит от коэффициентов. Распределение $f_Z(\boldsymbol{\zeta})$ позволяет получить оценку величины Z, связанную с ней неопределенность и доверительный интервал, содержащий истинное значение измеряемой величины с заданной доверительной вероятностью.

ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Во всех вышеприведенных рассуждениях и в дальнейшем изложении полагаем: латинские прописные буквы обозначают величины, латинские строчные – заданные или известные значения этих величин, а строчные греческие буквы – случайные переменные, связанные с возможными значениями величин. Распределения вероятностей обозначаются буквой f с индексами, указывающими на соответствующие величины. Выборочные распределения обозначаются буквой p .

Величина интерпретируется в соответствии с определением из словарей: «свойство явления, тела или вещества, которое может быть выражено количественно в виде числа с указанием отличительного признака как основы для сравнения» [2–5] (обычно – единиц измерения). В большинстве случаев векторы **X** и **Y** удовлетворяют этому определению. Ситуация с коэффициентами **A** не столь однозначна: хотя обычно их величины можно выразить числом и отличительным признаком, но не всегда возможно рассматривать их как «свойства явления, тела или вещества».

Например, если X – это сила, удлиняющая упругую пружину до длины Y, то коэффициент A_1 в функции $Y = A_1 + A_2 \cdot X$ представляет собой ненагруженную длину пружины, а коэффициент A_2 – ее жесткость, т. е. обратную величину коэффициента упругости. Аналогично у функции $Y = \exp(-A \cdot X)$ единственный коэффициент A может представлять константу распада радиоактивного вещества, «недораспавшаяся» доля которого через время X равна Y. В подобных случаях функция $F(X, \mathbf{A})$ определяется на основе физической модели, и коэффициенты действительно можно рассматривать как физические свойства.

Однако возможно, что явление, порождающее информацию о входных векторах, слишком сложно, неизвестно или недостаточно хорошо изучено. В этом случае вид функции $F(X, \mathbf{A})$ заранее не задан и должен быть выбран исследователем. Выбор является произвольным, хотя обычно определяется структурой оценок $\{x_i, y_i\}$ ($i = 1, 2, \ldots, m$), например, их видом на диаграмме в декартовых координатах. Такая ситуация типична при калибровке измерительного прибора, когда вектор **X** состоит из воздействий, задаваемых с помощью соответствующих мер, а вектор **Y** может включать реакции прибора на эти воздействия, разности между реакциями и воздействиями, их отношения и т. д. В этом случае коэффициенты **A** не могут рассматриваться как свойства калибруемого прибора, так как, вообще говоря, можно построить множество различных кривых, проходящих «как можно ближе» к оценкам входных координат, и каждая из этих кривых будет иметь собственные коэффициенты.

В любом случае, можно ли приписать коэффициентам физический смысл или нет, понимается, что они определяются оценочной моделью $G(\mathbf{X}, \mathbf{Y}, \mathbf{A}) = 0$, которая получается комбинацией критерия аппроксимации и заданной или выбранной формы функции $F(X, \mathbf{A})$. Коэффициенты при этом вынуждены принимать значения $\boldsymbol{\alpha}$, определяемые значениями $\boldsymbol{\xi}$ для $\mathbf{X}$ и $\boldsymbol{\eta}$ для $\mathbf{Y}$, независимо от того, что истинные значения координат остаются неизвестными. Как только модель установлена, она остается в силе вне зависимости от того, существуют ли коэффициенты как строго определенные свойства явления, тела или вещества, или же они были «придуманы» для учета неполного знания о фактической связи между координатами [6, 7].

В разделе 2 данной работы описывается процедура определения вышеупомянутых распределений вероятности. Важно подчеркнуть, что эти распределения рационально выводятся из всей доступной информации о рассматриваемых величинах. Эта информация включает экспериментальные данные, вероятностные модели процесса генерации данных, функциональные зависимости, с помощью которых одни величины выражаются через другие, а также любое априорное знание о величинах [8–11]. Более общий подход, сформулированный в [9, 12, 13], опирается на метод Монте-Карло (ММК) для численного нахождения интересующих распределений вероятности, который используется в этой работе.

ПРОЦЕДУРА АНАЛИЗА РАСПРЕДЕЛЕНИЙ

Первая цель анализа заключается в получении совместного распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$, которое описывает поведение входных координат $\mathbf{X}$ и $\mathbf{Y}$. При желании точечные оценки x_i и y_i можно интерпретировать как средние значения распределений компонентов векторов $f_{X_i}(\xi_i)$ и $f_{Y_i}(\eta_i)$ соответственно. Однако эти оценки не являются обязательными, поскольку распределение $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ нужно только для нахождения распределения $f_{\mathbf{A}}(\boldsymbol{\alpha})$, что является второй целью анализа. Третья возможная цель – получение распределения $f_Z(\zeta)$ некоторой величины Z, которая является функцией одного или нескольких коэффициентов $\mathbf{A}$, а возможно, и других величин. В данном разделе представлен общий план анализа; он легко расширяется на более высокие размерности, например, на построение и использование аппроксимирующих поверхностей [12–16]. Кратко также обсуждается связь предлагаемого метода с традиционной процедурой метода наименьших квадратов (МНК).

ПЕРВЫЙ ЭТАП: ПОЛУЧЕНИЕ РАСПРЕДЕЛЕНИЯ $F_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$

Получение распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ является первым этапом предлагаемой процедуры. Эта задача может быть простой или достаточно сложной. Ее нельзя сформулировать в общем виде, так как она сильно зависит от того, как определяются входные координаты $\mathbf{X}$ и $\mathbf{Y}$, каким образом получаются их оценки и являются ли они независимыми или нет. Ситуация, когда все координаты зависят друг от друга, может оказаться весьма трудной для анализа, но, к счастью, она встречается редко. Например, хотя изменение длины Y линейной пружины, очевидно, зависит от приложенной силы X, эта зависимость как раз моделируется функцией $Y = A_1+A_2 \cdot X$, подлежащей нахождению. Однако на первом этапе исследователя интересует не эта функциональная зависимость, а возможные отношения между информацией, имеющейся о векторах $\mathbf{X}$ и $\mathbf{Y}$. Если информацию о каждом из векторов можно собрать независимо, то их совместные распределения также будут независимы, и в этом случае $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta}) = f_{\mathbf{X}}(\boldsymbol{\xi})f_{\mathbf{Y}}(\boldsymbol{\eta})$. Независимость этих векторов в общем случае нельзя принимать как данность. Например, если X – сила, удлиняющая пружину до длины Y, то входные координаты $\mathbf{X}$ могут быть представлены массами гирь, подвешиваемых к свободному концу пружины (вертикальной). Если эти массы определяются путем сравнительных взвешиваний с эталонной гирей, они становятся сильно коррелированными. Аналогично, если удлинение пружины измеряется с помощью тензометра, то как неизвестная с достаточной точностью поправка на калибровку прибора, так и измерение ненагруженной длины влияют на все измерения $\mathbf{Y}$, так что они также оказываются коррелированными.

Тем не менее для упрощения часто предполагается, что влиянием этих корреляций можно пренебречь. Такое упрощение позволяет рассматривать любую входную величину как независимую от остальных, так что $f_{\mathbf{X}}(\boldsymbol{\xi}) = \prod_{i=1}^{m} f_{X_i}(\xi_i)$ и $f_{\mathbf{Y}}(\boldsymbol{\eta}) = \prod_{i=1}^{m} f_{Y_i}(\eta_i)$.

Если входные координаты все независимы, но для одной из них, скажем X_i, нет данных, то ее распределение должно быть либо задано как результат предыдущего анализа, либо о ней должна быть доступна достаточная информация для построения распределения, например, с использованием принципа максимальной энтропии. Например, если известно лишь, что истинное значение X_i находится в пределах заданного интервала, то $f_{X_i}(\xi_i)$ считается равномерным распределением на этом интервале.

Для имеющихся входных данных обычно предполагается, что они получены из выборочных распределений, которые в общем случае зависят от различных параметров. Типичная ситуация возникает, когда есть измеренные значения величины $\mathbf{x}_i = \{x_{i1},\ldots,x_{in_i}\}$, относящиеся к координате X_i,

каждое из которых считается независимо сгенерированным некоторым случайным процессом. Например, если процесс является гауссовским, то выборочное распределение имеет вид:

$$p(\chi_i|\xi_i,\sigma)=\prod_{j=1}^{n_i}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\xi_i-\chi_{ij})^2}{2\sigma_i^2}\right], \quad (1)$$

где ξ_i и σ_i обозначают значения X_i и стандартного отклонения S_i соответственно. Если имеется лишь одно наблюдение χ_i, то вполне естественно полагать его нормально распределенным, но в этом случае стандартное отклонение должно быть известно совершенно точно, чтобы можно было делать выводы относительно X_i.

Нормальный процесс является наиболее распространенной вероятностной моделью для описания изменчивости данных измерений, но не единственной. Другой пример – пуассоновский процесс, обычно применяемый, когда X_i есть интенсивность наступления некоторого события, а данные представляют собой серию подсчетов числа наступлений события за непересекающиеся интервалы времени. В этом случае выборочное распределение имеет вид:

$$p(\chi_i|\xi_i)=\frac{(\xi_i t)^{\chi_i}\exp(-\xi_i t)}{\prod_{j=1}^{n_i}\chi_{ij}!}, \quad (2)$$

где $\xi_i > 0$;

$$\chi_i=\sum_{j=1}^{n_i}\chi_{ij};$$

t – общее время наблюдения.

Для всех неточно известных параметров в выборочных распределениях требуются «априорные» функции. Эти априорные распределения могут быть заданы как известные распределения или построены на основе имеющейся информации. Однако встречаются ситуации, когда априори отсутствует какая-либо информация о некоторых или обо всех параметрах. В таких случаях следует использовать неинформативные априорные распределения. Выбор подобных распределений является спорным вопросом [17]; в настоящее время общепринята практика их получения по формальным правилам, основанным на структуре выборочных распределений [18].

Например, если параметры принадлежат семейству распределений типа «сдвиг – масштаб», как в (1), и при этом нет априорной информации ни о X_i, ни о S_i, соответствующее неинформативное априорное распределение имеет вид: $f_{X_iS_i}(\xi_i,\sigma_i)=1/\sigma_i$ [19]. Если выборочное распределение задано уравнением (2) и отсутствует априорная информация о X_i, используется априорное распределение: $f_{X_iS_i}(\xi_i)=1/\sqrt{\xi_i}$ [19].

Как показывают эти примеры, очень часто неинформативные априорные распределения оказываются функциями, не подлежащими нормировке. Все априорные распределения обновляются путем их умножения на соответствующие выборочные распределения с последующей нормализацией. Эта операция, основанная на теореме Байеса, приводит к «апостериорным» совместным распределениям неизвестных параметров. Например, умножение выборочного распределения (1) на априорное $1/\sigma_i$ дает совместное апостериорное распределение $f_{X_iS_i}(\xi_i,\sigma_i|\chi_{ij})$, распределение $f_{X_i}(\xi_i|\chi_{ij})$ отдельных компонент которого является t-распределением Стьюдента с параметрами: сдвига $\chi_i=\frac{1}{n_i}\sum_{j=1}^{n_i}\chi_{ij}$, масштаба $s_i=1/\sqrt{n_i}$ и числом степеней свободы $n_i - 1$, где $s_i^2=\frac{1}{n_i-1}\sum_{j=1}^{n_i}(\chi_{ij}-\chi_i)^2$. Частное распределение $f_{S_i}(\sigma_i|\chi_i)$, как правило, не представляет интереса при решении задачи аппроксимации, но можно показать, что оно является масштабированным обратным хи-квадрат распределением с $n_i - 1$ степенями свободы и средним значением: $(n_i-1)s_i^2/(n_i-3)$. Аналогично умножение выборочного распределения (2) на априорное распределение $1/\sqrt{\xi_i}$ дает для X_i апостериорное гамма-распределение с параметром формы $\chi_i+1/2$ и параметром нормировки t, среднее значение которого равно $(\chi_i+1/2)/t$. Эти результаты хорошо известны [20].

ВТОРОЙ ЭТАП: НАХОЖДЕНИЕ РАСПРЕДЕЛЕНИЯ $F_A(\alpha)$

После того как форма функции $Y=F(X, \mathbf{a})$ определена (на основе физического смысла рассматриваемой задачи или простого анализа данных), проблема нахождения распределения $f_A(\alpha)$ на втором этапе не зависит от того, каким образом было получено распределение $f_{XY}(\xi,\eta)$. Поэтому для нее можно предложить достаточно общий подход.

Как упоминалось выше, эта задача во многом определяется критерием, используемым для аппроксимации функции по оценкам координат. Наиболее распространенным критерием является МНК, который заключается в решении, как правило, неявной системы:

$$G(\mathbf{X}, \mathbf{Y}, \mathbf{A}) = 0, \tag{3}$$

возникающей из минимизации взвешенной суммы квадратов остатков $R_i = Y_i - F(X_i, \mathbf{A})$:

$$\sum_{j=1}^{n_i} w_i R_i^2 \to \min\ , \tag{4}$$

где весовые коэффициенты w_i (все могут быть равны единице) выбираются в соответствии с определенным правилом. Следует отметить, что этот критерий аппроксимации является частным случаем так называемой L_q-минимизации [21, 22], когда L_q-норма является суммой модулей остатков, возведенных в степень $q \geq 1$.

Система (3) является моделью для процедуры оценки в том смысле, что для заданных значений $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ (для векторов **X** и **Y**) значения оценок $\boldsymbol{\alpha}$ (для вектора **A**) оказываются фиксированными. Обычно исходных данных или априорной информации о коэффициентах **A** нет. В этом случае их совместное распределение получается аналитически путем преобразования распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ с последующим разделением его на распределение отдельных компонент вектора или численно путем решения уравнения (3). В некоторых случаях могут существовать исходные данные или априорные сведения о коэффициентах **A**, но этот случай выходит за рамки настоящей работы.

Как отмечалось выше, в данном подходе оценки $\mathbf{x} = (x_1,\ldots,x_m)^T$ и $\mathbf{y} = (y_1,\ldots,y_m)^T$ могут быть заданы заранее или вычислены впоследствии как средние значения распределений отдельных компонентов $f_{X_i}(\xi_i)$ и $f_{Y_i}(\eta_i)$ ($i = 1,\ldots,m$). Однако эти данные не являются строго обязательными – требуется лишь наличие информации для получения совместного распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$. Аналогично оценки $\mathbf{a} = (a_1,\ldots,a_n)^T$ коэффициентов могут быть вычислены как средние значения распределений $f_{A_i}(\alpha_i)$ ($i = 1,\ldots,n$), но и это может оказаться не всегда необходимым.

ПРОВЕРКА СОГЛАСОВАННОСТИ

После получения распределения $f_{\mathbf{A}}(\boldsymbol{\alpha})$ следует проверить, согласованы ли остатки R_i с неопределенностями, содержащимися в распределениях $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ и $f_{\mathbf{A}}(\boldsymbol{\alpha})$. В предельном случае, если неопределенностями, связанными с оценками всех входных координат, можно пренебречь, распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ и $f_{\mathbf{A}}(\boldsymbol{\alpha})$ исчезают или, точнее, вырождаются в многомерные дельта-функции [8]. Значения коэффициентов тогда находятся путем минимизации суммы: $\sum_{i=1}^{m}(y_i - F(x_i, \mathbf{a}))^2$, что приводит к детерминированной системе уравнений $G(\mathbf{x}, \mathbf{y}, \mathbf{a}) = 0$. Из решения этой системы можно вычислить величины остатков $r_i = y_i - F(x_i, \mathbf{a})$. Здесь с ними не будет связано никакой неопределенности. Однако, если все входные координаты известны абсолютно точно, кривая не должна «подгоняться»: она обязана проходить через все точки, т. е. все остатки должны обращаться в ноль. Если этого не происходит, то кривая либо неадекватно описывает физику задачи, либо допущена ошибка в процессе получения данных, либо предположение о пренебрежимости неопределенностей противоречит выбранной форме кривой. Естественно, этот предельный случай интереса не представляет.

В общем случае остатки $R_i = Y_i - F(X_i, \mathbf{A})$ являются случайными величинами с совместным распределением $f_R(\rho)$, которое можно определить, выразив коэффициенты через входные координаты посредством модели (3) с последующим определением распределений $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ и $f_{\mathbf{A}}(\boldsymbol{\alpha})$. Согласованность можно установить, вычисляя площадь под распределениями отдельных компонентов $f_{R_i}(\rho_i)$, причем интегрирование надо проводить от нуля в противоположную сторону от расположения соответствующего среднего значения r_i. Данные считаются согласованными с функцией $Y = F(X, \mathbf{A})$, если все такие вероятности превышают некоторое произвольное малое значение, например 1 %. Так как распределения $f_{R_i}(\rho_i)$ часто являются нормальными (или близкими к ним), эту проверку можно упростить: достаточно убедиться, что все отношения $|r_i|/u(r_i)$ меньше некоторого числа, например 3, где $u(r_i)$ – стандартная неопределенность распределения $f_{R_i}(\rho_i)$.

ТРЕТИЙ ЭТАП: ПОЛУЧЕНИЕ РАСПРЕДЕЛЕНИЯ $F_Z(\boldsymbol{\zeta})$

Если измеряемыми величинами являются сами коэффициенты **A** (прямые измерения), тогда третий этап отсутствует. Например, может представлять интерес коэффициент A_2, соответствующий величине, обратной жесткости пружины, у функции $Y = A_1 + A_2 \cdot X$. В общем случае измеряемая величина Z определяется как функция одного или нескольких коэффициентов **A** и, возможно, других величин **W**. Тогда необходимо установить распределение $f_Z(\boldsymbol{\zeta})$. При этом известная информация

о величинах **W** должна быть задана известным распределением вероятности $f_{\mathbf{W}}(\boldsymbol{\omega})$, которое в большинстве случаев можно считать независимым от распределения $f_{\mathbf{A}}(\boldsymbol{\alpha})$. Распределение $f_Z(\zeta)$ затем получается путем распространения распределений $f_{\mathbf{A}}(\boldsymbol{\alpha})$ и $f_{\mathbf{W}}(\boldsymbol{\omega})$ через модель оценивания вида:

$$H(\mathbf{A}, \mathbf{W}, Z) = 0. \quad (5)$$

Типичная ситуация – когда измеряемая величина представляет собой воздействие X_0, соответствующий отклику Y_0 измерительного прибора. Если калибровочная функция взаимно-однозначна в окрестности этих координат, модель (5) сводится просто к $W - F(Z, \mathbf{A}) = 0$, где $W = Y_0$, а $Z = X_0$.

ЧИСЛЕННОЕ ВЫЧИСЛЕНИЕ РАСПРЕДЕЛЕНИЙ

Из приведенной общей схемы видно, что постановка задачи оценки неопределенности, связанной с построением и использованием аппроксимирующей кривой, относительно проста. Однако ее численная реализация может вызвать трудности. Действительно, редко удается получить распределение $f_{\mathbf{A}}(\boldsymbol{\alpha})$ на втором этапе с использованием, например, процедуры из [8], которая сочетает аналитические формулы с численными методами интегрирования. Кроме того, само распределение $f_{\mathbf{A}}(\boldsymbol{\alpha})$ может не быть конечной целью: оно может потребоваться на третьем этапе для вычисления распределения $f_Z(\zeta)$ измеряемой величины, определенной уравнением (5).

К счастью, распределения $f_{\mathbf{A}}(\boldsymbol{\alpha})$ и $f_Z(\zeta)$ могут быть численно аппроксимированы с использованием ММК [9, 23, 24], который в общих чертах можно описать следующей последовательностью действий: 1) сгенерировать случайные выборки $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ из распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$; 2) вычислить значения $\boldsymbol{\alpha}$, удовлетворяющие системе $G(\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\alpha}) = 0$; 3) сгенерировать случайную выборку $\boldsymbol{\omega}$, подчиняющуюся распределению $f_{\mathbf{W}}(\boldsymbol{\omega})$; 4) вычислить значение ζ, получающееся из уравнения $H(\boldsymbol{\alpha}, \boldsymbol{\omega}, \zeta) = 0$. Так как $\boldsymbol{\alpha}$ и ζ являются выборками, подчиняющимися распределениям $f_{\mathbf{A}}(\boldsymbol{\alpha})$ и $f_Z(\zeta)$ соответственно, эти распределения аппроксимируются повторением описанной последовательности действий большое число раз. Согласно [3, 9, 24], для получения статистически устойчивых результатов часто достаточно около миллиона повторных экспериментов. Однако, поскольку нет гарантии, что именно это или любое заранее заданное число будет достаточным, в [3, 9] описана процедура адаптивного выбора числа повторений. Дополнительные рекомендации по этому вопросу были опубликованы недавно [25].

СВЯЗЬ С ТРАДИЦИОННОЙ ПРОЦЕДУРОЙ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

Традиционный МНК отличается от описанного в подразделе 2.2. Классический МНК обычно называют *регрессионным анализом методом наименьших квадратов* [26, 27], и он основывается на интерпретации точек $\{x_i, y_i\}$ не как оценок входных величин, а как реализаций случайных величин с соответствующими выборочными распределениями. Переменные, связанные с данными **x**, называют обычно *независимыми*, а переменные, связанные с данными **y**, – *зависимыми* или *откликами* [26–29].

Традиционная регрессия МНК описана в огромном числе монографий и статей [26–32]. В этой трактовке в классическом МНК первый этап отсутствует, так как распределение $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$ здесь не используется. Вместо этого вероятностные модели, применяемые для построения выборочных распределений, напрямую включаются в выражение, подлежащее минимизации. В результате возникает множество вариантов, которые в зарубежной литературе классифицируются следующим образом: 1) *обыкновенный* МНК, который применяется, когда независимые переменные **x** известны точно, а отклики **y** содержат случайные ошибки, распределенные независимо и с одинаковыми дисперсиями; 2) *взвешенный* МНК исходит из тех же предположений, что и обыкновенный, за исключением допущения, что дисперсии откликов **y** могут быть различными; 3) МН *полных квадратов* (МНПК) применим, когда и **x**, и **y** известны с ошибками, но при этом все ошибки некоррелированные; 4) *обобщенный* МНК совпадает с МНПК, погрешности могут быть коррелированными. Ни одна из классификаций 1)–4) не соответствует методу, предложенному в данной работе, поскольку критерий (4) может быть использован всегда – вне зависимости от вероятностных моделей, примененных для определения распределения $f_{\mathbf{XY}}(\boldsymbol{\xi},\boldsymbol{\eta})$. Кроме того, важно подчеркнуть, что предлагаемый метод не является и *байесовской регрессией*, краткое описание которой дано, например, в [33].

ПРИМЕРЫ

В этом разделе приводятся три примера. Первые два касаются достаточно распространенных задач аппроксимации константой и прямой линией при использовании нормальных выборочных распределений. Третий пример иллюстрирует аппроксимацию нелинейной кривой по данным, подчиняющимся пуассоновскому закону распределения.

АППРОКСИМАЦИЯ КОНСТАНТОЙ

Оценка константы по экспериментальным данным – самый простой случай предлагаемой процедуры. Она применяется, например, когда некоторая величина A измеряется в условиях воспроизводимости [2], т. е. в разных местах, разными наблюдателями или с использованием различных измерительных систем. В этом случае координаты **X** отсутствуют, и функция $Y = F(X, \mathbf{A})$ принимает простой вид $Y = A$, а координаты **Y** представляют собой результаты прямых измерений величины A

в различных условиях. Примерами могут служить: определение значений фундаментальных физических констант, измерение свойств элементарных частиц или обработка ключевых сличений, выполняемых метрологическими службами различных организаций из различных стран.

В качестве конкретного примера рассмотрим ключевое сличение, получившее название в зарубежной литературе CCQM-K25 и описанное в работе [34], в рамках которого измерялись конгенеры полихлорированных бифенилов (ПХБ) в отложениях. Результаты для разновидности ПХБ 28 (2,4,4′-трихлорбифенил) приведены в таблице 1; они включают сами оценки y_i, стандартные неопределенности $u(y_i)$ и число степеней свободы ν_i для измерений, выполненных шестью разными лабораториями (m = 6). В итоговом отчете указано, что ключевое значение сличения (*KCRV*) равно $\bar{y}$ = 33,64 нг/г, его стандартная неопределенность $u(y)$ = 0,60 нг/г, но при этом не объяснено, как были вычислены эти результаты. При этом данные результаты совпадают с результатами, полученными согласно классическим соотношениям [26–29, 35]: $y = \frac{1}{m}\sum_{i=1}^{m} y_i$ и $u(y) = \frac{1}{m(m-1)}\sum_{j=1}^{m}(y_i - \bar{y})^2$, из чего можно заключить, что при вычислениях полностью были проигнорированы исходные стандартные неопределенности и число степеней свободы.

Таблица 1 – Оценки y, стандартные неопределенности $u(y)$ и степени свободы ν массовой доли ПХБ 28, представленные лабораториями, участвующими в ключевом сравнении CCQM-K25 из работы [34]

№ лаборатории	y, нг/г	$u(y)$, нг/г	ν
1	34,3	1,03	60
2	32,9	0,69	4
3	34,53	0,83	18
4	32,42	0,29	2
5	31,9	0,4	13
6	35,8	0,38	60

В статье [36] данные этого ключевого сличения анализировались с использованием различных статистических моделей. Представляемый ниже в этой работе анализ существенно отличается от описанных в [34, 36]. На первом этапе будем полагать, что исходные результаты лабораторий (не приведенные явно в [34]) являются независимыми случайными величинами, распределенными по нормальному закону с неизвестными средними Y_i и неизвестными стандартными отклонениями S_i; кроме того, измерения y_i из таблицы 1 являются средними значениями, полученными из $n_i = m_i + 1$ измерений, а стандартные неопределенности рассчитаны из выражения $u(y) = s_i / \sqrt{n_i}$, где s_i^2 – выборочные дисперсии исходных экспериментальных данных. Соответственно, распределения $f_{Y_i}(\eta_i)$ будут рассматриваться как t-распределения с параметрами: среднее значение y_i, масштабный коэффициент $u(y_i)$ и число степеней свободы ν_i.

Интерес представляет вопрос, можно ли считать эти распределения независимыми. В данном ключевом сличении CCQM-K25, вероятно, существовали корреляции, вызванные, например, использованием одних и тех же эталонных материалов. Однако лаборатории не могли учесть эти корреляции, так как, предположительно, предоставляли свои результаты независимо – в полном неведении о действиях других участников процесса. Таким образом, ковариации могли быть учтены только лабораторией, обрабатывающей информацию, полученную от всех участников. Но для этого потребовались бы гораздо более детальные данные, чем просто тройки $\{y_i, u(y_i), \nu_i\}$ из таблицы 1, приведенные в отчете.

Теперь рассмотрим второй этап. Минимизация функции невязки: $\sum_{j=1}^{m}(Y_i - A)^2 \to \min$ приводит к модели оценки:

$$A = \frac{1}{m}\sum_{i=1}^{m} Y_i, \tag{6}$$

из которой распределение $f_A(\alpha)$ получается из распределений $f_{Y_i}(\eta_i)$. Однако следует отметить, что поскольку лаборатория № 4 выполнила всего три измерения, разброс значений, выборки которых берутся из t-распределения, назначенного для этой лаборатории, оказывается довольно большим. Из математической статистики следует [9], что в этом случае снятие трех миллионов отсчетов позволяет достичь достаточно устойчивых статистических результатов. Среднее значение и стандартное отклонение распределения $f_A(\alpha)$ составили: a = 33,64 нг/г и $u(a)$ = 0,35 нг/г соответственно, что совпадает со значением *KCRV*, приведенным в [21].

В модели (6) все весовые коэффициенты были выбраны равными единице. Чаще их выбирают как обратные величины дисперсий распределений $f_{Y_i}(\eta_i)$. При таком подходе модель оценки получается следующей [9, 37]: $A' = \sum_{i=1}^{m} \frac{Y_i}{w_i} \Big/ \left(\sum_{i=1}^{m} \frac{1}{w_i} \right)$, где $w_i = u_i^2 \nu_i / (\nu_i - 2)$. Заметим, что эта процедура фактически означает исключение результатов лаборатории № 4. Вновь при проведении $3 \cdot 10^6$ отсчетов были получены: $a' = 34{,}05$ нг/г, $u(a') = 0{,}26$ нг/г соответственно. Эти значения отличаются от a и $u(a)$, поскольку модели оценок, определяющие величины A и A', различны.

Для проверки согласованности результатов, представленных лабораториями, необходимо проанализировать неопределенность остаточных членов $R_i = Y_i - A$ или $R'_i = Y_i - A'$. В контексте ключевых сличений эти остатки называются *степенями эквивалентности* [35, 38]. Чтобы построить распределения остатков, можно провести повторные выборки η_i и α из распределений $f_{Y_i}(\eta_i)$ и $f_A(\alpha)$ или $f_{A'}(\alpha')$ и вычислить разности между соответствующими отсчетами этих выборок. Однако такой подход игнорировал бы тот факт, что оба коэффициента A и A' коррелированы с входными величинами. Более корректной процедурой является подстановка определения этих коэффициентов в выражения для остатков и проведения выборки исключительно из распределений $f_{Y_i}(\eta_i)$. Разумеется, это можно делать одновременно с вычислением $f_A(\alpha)$ или $f_{A'}(\alpha')$.

В таблице 2 приведены оценки и стандартные неопределенности для $R_i = Y_i - A$ и $R'_i = Y_i - A'$. Они показывают, что значения, полученные лабораториями № 5 и № 6, не согласованы с результатами остальных, поскольку при обоих критериях наименьших квадратов оценки их остатков отличаются от нуля более чем на три стандартные неопределенности. Однако может оказаться, что именно одна из этих двух лабораторий предоставила оценку измеряемой величины, наиболее близкую к ее истинному значению. Ни один статистический тест не может однозначно утверждать, так это или нет.

Таблица 2 – Оценки и стандартные неопределенности остатков $R_i = Y_i - A$ и $R'_i = Y_i - A'$ в ключевом сравнении CCQM-K25. Все результаты даны в нг/г

№ лаборатории	r	$u(t)$	r'	$u(r')$
1	0,66	0,92	0,25	1,02
2	–0,74	0,86	–1,15	0,94
3	0,89	0,79	0,48	0,84
4	–1,22	0,85	–1,63	0,99
5	–1,74	0,49	–2,15	0,35
6	2,16	0,46	1,75	0,29

АППРОКСИМАЦИЯ ПРЯМОЙ ЛИНИЕЙ

Аппроксимация прямой линией некоторого набора точек на плоскости – очень распространенная задача в метрологии. Для иллюстрации процедуры, предложенной в настоящей статье, рассмотрим пример Н.3 из [1, 39]. Проводится калибровка термометра путем получения $m = 11$ точек $\{x_i, y_i\}$, где измерения x_i – это показания температуры $t_i - 20$ °C, а y_i – поправки к показаниям термометра, т. е. разности между опорными температурами t_{Ri} и соответствующими показаниями t_i. Эти данные даны в таблице 3.

В [1, 39] данный пример сформулирован довольно противоречиво. Так, в пункте Н.3.1 утверждается, что опорные температуры известны. Это можно интерпретировать как отсутствие неопределенности, связанной с их значениями, что следует из таблицы 3. Однако далее говорится, что показания t_i имеют пренебрежимо малую неопределенность. Если бы это было так, аппроксимация прямой линией данных была бы невозможной: ордината, вычисленная по линии при абсциссе x_i, отличалась бы от соответствующего значения y_i, которое предполагается известным точно. Иными словами, аппроксимирующая линия оказалась бы несогласованной с заявленными допущениями. В то же время утверждается, что поправки $y_i = t_{Ri} - t_i$ и температуры t_i являются результатами измерений. Из этого следует, что ни одна из этих величин не является идеально известной.

Эти вопросы были подняты в [6], где пример Н.3 анализировался с использованием традиционного МНК. Анализ, предлагаемый авторами, отличается от [6]. Будем считать, что неопределенности, связанные с опорными температурами t_{Ri}, пренебрежимо малы (как обычно и бывает в калибровке), а температуры t_i представляют собой оценки не идеально известных величин T_i. Тогда независимыми величинами, участвующими в калибровке, будут: $X_i = T_i - 20$ и $Yi = t_{Ri} - T_i$. На первом этапе будем предполагать, что в процессе калибровки распределения $f_{T_i}(\tau_i)$ являются равномерными на интервалах $(t_i - e,\ t_i + e)$, где e – абсолютное значение максимальной погрешности всех показаний, известное из опыта работы с данным типом термометров и равное 0,01 °C.

Таблица 3 – Данные калибровки термометра. Температуры t – отклики прибора на известные опорные температуры t_R. Значения r – оценки остатков для корректировки, иллюстрируемой рисунком 1, u(r) – соответствующие стандартные неопределенности. Все значения указаны в градусах Цельсия

№ опыта	$x = t–20$	$y = t_R–t$	r	$u(r)$
1	1,521	–0,171	–0,003	0,003
2	2,012	–0,169	–0,002	0,003
3	2,512	–0,166	–0,0003	0,002
4	3,003	–0,159	0,006	0,002
5	3,507	–0,164	–0,0005	0,002
6	3,999	–0,165	–0,003	0,002
7	4,513	–0,156	0,005	0,002
8	5,002	–0,157	0,003	0,002
9	5,503	–0,159	0,0002	0,002
10	6,010	–0,161	–0,003	0,003
11	6,511	–0,160	–0,003	0,003

На втором этапе выбранная функция невязки, используемая при калибровке, будет $\sum_{i=1}^{m}(Y_i - A_1 - A_2 X_i)^2$. Ее минимизация с использованием классического МНК дает оценки коэффициентов [26–29, 37]: $\mathbf{A} = \left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{Y}$, где $\mathbf{A} = (A_1, A_2)^T$; $Y = (Y_1,\ldots,Y_{11})^T$; $\mathbf{G}$ – матрица плана эксперимента размера $2\times m$, в первом столбце которой стоят единицы, а во втором – величины X_i ($i = 1,\ldots,11$). В результате 10^6 опытов ММК были получены оценки: $a_1 = –0{,}1712$ °C и $a_2 = 0{,}0022$ со стандартными неопределенностями $u(a_1) = 0{,}0048$ °C и $u(a_2) = 0{,}0011$ и коэффициентом корреляции $r(a_1, a_2) = –0{,}931$. Оценки остаточных величин и их неопределенностей приведены в таблице 3. Отношения этих значений во всех 11 экспериментах меньше 3, поэтому аппроксимирующая прямая (показанная на рисунке 1) согласуется с неопределенностями в исходных данных.

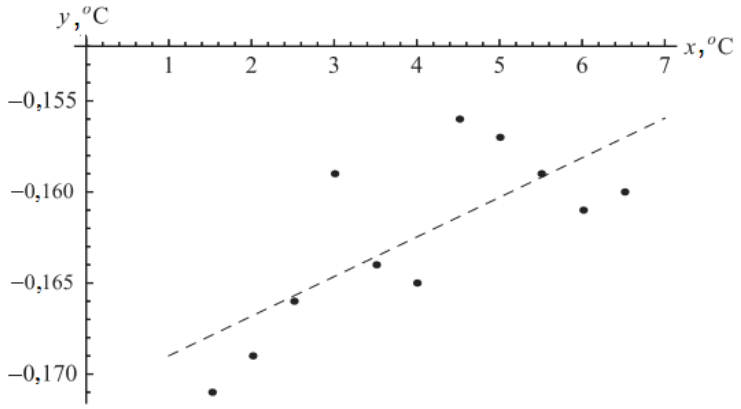


Рисунок 1 – Аппроксимация прямой для аддитивных поправок $y = t_R – t$, полученных из показаний термометра t. Координаты $x = t – 20$ °C

При рассмотрении третьего этапа предположили, что среднее значение трех показаний термометра составило $t_0 = 30$ °C, а их выборочное стандартное отклонение равно $s_0 = 17$ мК. Эта температура лежит вне диапазона калибровки, поэтому, чтобы получить распределение, связанное с откорректированной температурой T_c, необходимо предположить дополнительно, что калибровочная прямая может быть использована и в этой точке. Соответственно, 10^6 значений τ_0 были сгенерированы подчиняющимися t-распределению с параметрами: среднее значение t_0, масштабный коэффициент $s_0/\sqrt{3}$ и двумя степенями свободы. Для каждого значения вычислялось $\tau_c = \tau_0 + \alpha_1 + \alpha_2(\tau_0 – 20)$, где α_1 и α_2 были взяты из сохраненных выборок распределения $f_A(\alpha)$. Таким образом, неопределенности, связанные с коэффициентами **A**, были учтены. (Здесь максимальная ошибка показаний e не использовалась, так как условия эксплуатации отличаются от условий калибровки.)

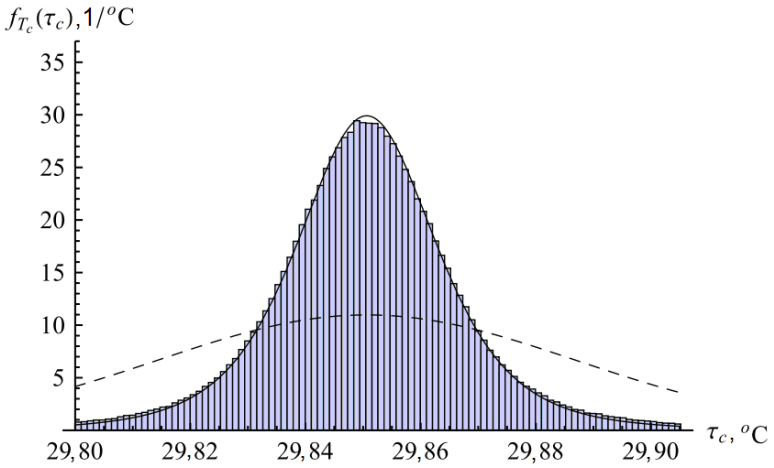


Рисунок 2 – Нормализованная гистограмма, аппроксимирующая распределение $f_{T_c}(\tau_c)$ откорректированной температуры, соответствующей показаниям t_0 = 30°C. Среднее значение равно t_c = 29,85 °C, стандартное отклонение – $u(t_c)$ = 0,036 °C. Сплошная кривая – t-распределение с параметрами: t_c, масштабным коэффициентом $u(t_c)/2{,}9$ и 4 степенями свободы. Пунктирная кривая – нормальное распределение с тем же средним t_c и стандартным отклонением $u(t_c)$

Среднее значение величин τ_c равно $t_c = 29{,}85$ °C, а их стандартное отклонение – $u(t_c) = 0{,}036$ °C. Как видно из рисунка 2, распределение $f_{T_c}(\tau_c)$ является симметричным, но не нормальным. Оно близко к *t*-распределению с параметрами: t_c, масштабным коэффициентом $u(t_c)/2{,}9$ и числом степеней свободы 4. Симметричный доверительный 95%-й интервал, вычисленный согласно [9], равен (29,81; 29,89) °C. Следует подчеркнуть, что доверительная вероятность 0,95, связанная с этим интервалом, носит условный характер и зависит от всех сделанных допущений.

АППРОКСИМАЦИЯ НЕЛИНЕЙНОЙ КРИВОЙ

Предположим, что имеется детектор радиоактивных частиц, который калибруется путем независимого подсчета числа частиц, испускаемых семью образцами радиоактивных эталонных материалов с очень большим периодом полураспада. Массы всех образцов одинаковы, а их аттестованные активности xi находятся в диапазоне от 1 до 50 Бк. Эти значения приведены в таблице 4 вместе с соответствующими стандартными неопределенностями u(xi). Также в ней указаны числа частиц Ni, зарегистрированных детектором за идеально известные промежутки времени ti. Наилучшей оценкой эффективности установки E является e = 0,300 со стандартной неопределенностью u(e) = 0,005. В работе [40] показано, что эффективность E входит в соответствующие выражения только как мультипликативный коэффициент у времени измерения t. Уровень фона также был измерен в процессе калибровки и оказался пренебрежимо малым. Диаграмма рассеяния значений xi и yi = Ni/(eti) показывает, что детектор частиц не является идеально линейным (рис. 3).

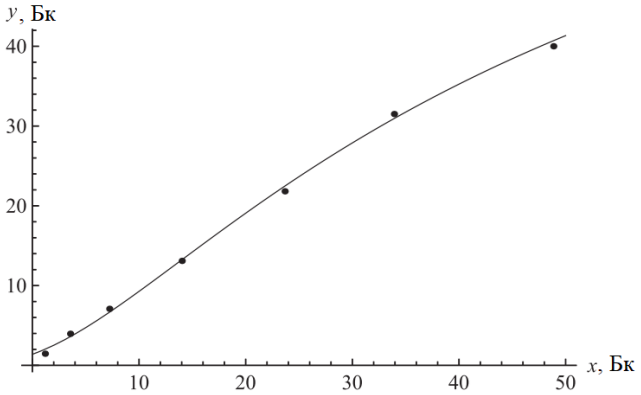


Рисунок 3 – Калибровочная кривая детектора частиц с использованием средних значений оценок a из таблицы 5 в качестве оценок коэффициентов A. Точки соответствуют координатам точек данных xi и yi = Ni/(eti)

Такое поведение можно всегда описать математически, выбрав в качестве калибровочной функции полином некоторой степени [7, 27, 29]. Чтобы продемонстрировать возможности предлагаемой методики вместо полиномов в работе выбрана аппроксимирующая кривая: $Y = A_1 \exp\left(\frac{A_2}{A_3 + X}\right)$, где X – активность образца, а Y – соответствующая измеренная активность.

Таблица 4 – Калибровочные данные для детектора частиц (использовались образцы с аттестованными активностями x и соответствующими стандартными неопределенностями u(x). Также приведено число частиц N, зарегистрированное детектором за время t)

№ образца	x, Бк	u(x), Бк	N	t, c
1	1,178	0,015	105	240
2	3,542	0,040	234	197
3	7,213	0,079	340	160
4	14,01	0,17	550	140
5	23,66	0,28	720	110
6	33,91	0,33	690	73
7	48,85	0,51	480	40

Использовался ММК путем генерации значений ξi, нормально распределенных со средними значениями xi и стандартными неопределенностями u(xi). В соответствии с изложенным в подразделе 2.1 значения ηi генерировались подчиняющимися гамма-распределениям с параметрами формы Ni+1/2 и коэффициентами нормировки λiti. Так как отношение e/u(e) превышает три, значения эффективности λi генерировались нормальными со средним e и стандартным отклонением u(e) [26–29, 37]. Далее вычислялись тройки коэффициентов $\alpha = (\alpha 1, \alpha 2, \alpha 3)T$, доставляющих минимум функции невязки:

$$\sum_{i=1}^{m}\left(\eta_i - \alpha_1 \exp\left(\frac{\alpha_2}{\alpha_3 + \xi_i}\right)\right)^2 \to \min . \quad (7)$$

Повторяя опыты ММК 106 раз, было оценено распределение $f_A(\alpha)$. Средние значения, стандартные отклонения и корреляционная матрица [14–16] коэффициентов A приведены в таблице 5. На рисунке 3 показана аппроксимирующая кривая вместе с точками данных. Все отношения остаточных значений ri/u(ri), вычисленные аналогично примеру с прямой линией, оказались меньше 3, что свидетельствует о согласованности выбранной калибровочной функции с данными. Однако полученная кривая не проходит через начало координат. Это указывает на то, что выбор функции был не самым удачным. В частности, ее не следует использовать для предсказания активностей менее 1–2 Бк.

Предположим теперь, что измеряется образец с неизвестной активностью X0, и детектор зарегистрировал N0 = 328 импульсов за время t0 = 170 с в условиях, когда нельзя пренебрегать фоновой активностью. Тогда дополнительно измеряется контрольный образец, давший Ne = 27 импульсов за время te = 350 с. Оба образца идентичны по массе и форме образцам, использованным при калибровке. Для вычисления активности исследуемого образца было сгенерировано 106 значений η0 и ηe из соответствующих гамма-распределений, а также отобраны тройки коэффициентов a из сохраненных выборок распределения $f_{\mathbf{A}}(\alpha)$. Соответствующие значения ξ найдено из решения уравнения $\eta_0 - \eta_e = \alpha_1 \exp\left(\frac{\alpha_2}{\alpha_3 + \xi_i}\right)$. Оказалось, что ни одно из значений разности $\eta_0 - \eta_e$ не было отрицательным. На рисунке 4 представлено численное приближение распределения $f_{X_0}(\xi)$ активности образца. Среднее значение составило x0 = 6,63 Бк, стандартное отклонение – u(x0) = 0,48 Бк.

Из рисунка 4 видно, что распределение $f_{X_0}(\xi)$ хорошо аппроксимируется нормальным распределением, что упрощает аналитическое вычисление доверительного интервала, если это потребуется. Снова подчеркнем, что вероятность, связанная с любым таким интервалом, является условной в зависимости от всех сделанных допущений.

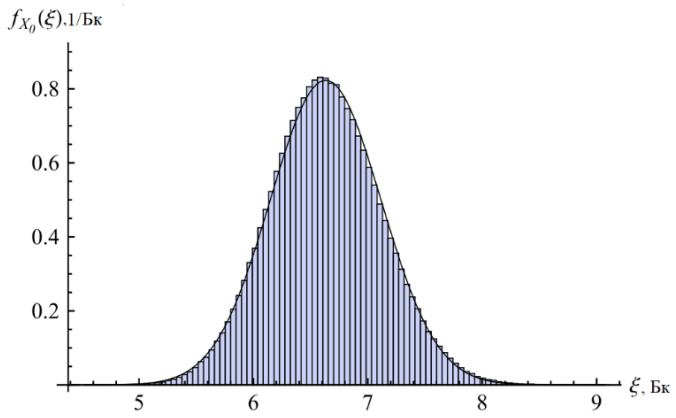


Рисунок 4 – Нормализованная гистограмма, полученная ММК для аппроксимации распределения $f_{X_0}(\xi)$ активности измеряемого образца. Сделано N0 = 328 отсчетов за t0 = 170 с. Также измеряется контрольный образец, давший Ne = 27 отсчетов за te = 350 с. Кривая соответствует нормальному распределению с центром в ξ = 6,63 Бк и стандартным отклонением 0,48 Бк

Следует отметить, что выбранная калибровочная функция является нелинейной относительно коэффициентов A, а статистические модели координат Y не являются нормальными случайными величинами. Но этот факт не создает никаких трудностей для предлагаемой методики, кроме необходимости численной минимизации выражения (7).

Таблица 5 – Средние значения, стандартные отклонения и корреляционная матрица коэффициентов A

Коэффициент	Среднее значение, Бк	Стандартное отклонение, Бк	Корреляционная матрица		
			A1	A2	A3
A1	94,9	17,7	1,0	–0,984	0,956
A2	–51,8	11,0	–0,984	1,0	–0,991
A3	12,3	2,7	0,956	–0,991	1,0

Графическое представление (гистограмма, плотность распределения) является удобным инструментом для визуальной оценки характера распределения результата. В ряде случаев (например, при логнормальном или треугольном распределении входов) распределение результата может быть существенно асимметричным, и в таких ситуациях метод Монте-Карло дает более реалистичную картину неопределенности по сравнению с классическим линейным подходом [9, 10].

ЗАКЛЮЧЕНИЕ

Представлена процедура получения совместного распределения вероятностей коэффициентов кривой A, аппроксимирующей два вектора входных данных X и Y. Процедура начинается с определения совместного распределения, связанного с этими векторами, таким образом, чтобы была учтена вся доступная информация о величинах, которые они представляют. Далее строится математическая модель измерения, связывающая коэффициенты с координатами. Модель может быть получена путем минимизации, в общем случае, взвешенной суммы квадратов разностей между значениями координат компонентов одного вектора и координатами элементов второго вектора, предсказанными аппроксимирующей кривой. При этом выбор функции невязки остается за исследователем и может быть отличным от используемого в данной работе. Распределение для входных координат, преобразованных принятой моделью, определяет распределение для коэффициентов модели (наличия данных или априорной информации, о которых не предполагается). Процедуру удобно проводить с помощью ММК, описанного в [9, 35]. Наконец, может быть определена величина Z, зависящая от одного или нескольких коэффициентов аппроксимирующей кривой и, возможно, от других величин, в соответствии с дальнейшей математической моделью измерения. Распределение вероятностей для Z также может быть получено с помощью процедуры ММК.

Получаемые распределения вероятностей $f_{\mathrm{A}}(\alpha)$ и $f_Z(\zeta)$ имеют во многом субъективный характер, так как зависят от нескольких сделанных предположений и выбора, проводимого исследователем, например, формы самой аппроксимирующей кривой. Иногда эта кривая предопределена физическим смыслом задачи, задающей входные координаты. В других случаях, например, при калибровке допустимы самые разнообразные кривые. В любом случае следует провести проверку, чтобы убедиться, что выбранная кривая согласуется с данными и априорной информацией.

Наиболее распространенным критерием качества аппроксимации является МНК (в работе используется модифицированная версия), однако предлагаемая методика допускает рассмотрение и других

критериев (функций невязки). Наконец, модели выборки, используемые для получения распределения fXY(ξ,η), должны выбираться исходя из предполагаемых характеристик процесса генерации данных. Следовательно, информация, представленная распределениями $f_A(\alpha)$ и $f_Z(\zeta)$, должна интерпретироваться как условная, с учетом всех этих выборов и предположений, сделанных исследователем.

Список источников

1. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement. GUM 1995 with Minor Corrections, Joint Committee for Guides in Metrology. – JCGM 100, 2008. – URL: http://www.bipm.org/ utils/common/documents/jcgm/JCGM_100_2008_E.pdf (дата обращения: 29.08.25).

2. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM), Joint Committee for Guides in Metrology. – JCGM 200, 2008. – URL: http://www.bipm.org/utils/common/documents/jcgm/ JCGM_200_2008.pdf> (дата обращения: 29.08.25).

3. Международный словарь по метрологии. Основные и общие понятия и соответствующие термины. – Санкт-Петербург : НПО «Профессионал», 2010. – 81 с.

4. Балабан, О. М. Измерения длины в Древней Греции: неопределенность стандартов в золотой век Олимпийских игр / О. М. Балабан, В. В. Балабан, А. А. Львов, М. С. Светлов // Проблемы управления, обработки и передачи информации : сб. тр. VI Междунар. науч. конф. – Саратов : ООО СОП «Лоди», 2019. – С. 611–620.

5. Шаронов, П. А. Альтернативный подход к выражению неопределенности измерения в эксперименте / П. А. Шаронов, Е. Г. Умнова, Н. С. Вагарина, А. А. Львов, М. С. Светлов // Математическое моделирование, компьютерный и натурный эксперимент в естественных науках. – 2021. – № 3; URL: mathmod.esrae.ru/35-130 (дата обращения: 27.08.2025). DOI: 10.24412/2541-9269-2021-3-10-26.

6. Willink, R. Estimation and Uncertainty in Fitting Straight Lines to Data: Different Techniques / R. Willink // Metrologia. – 2008. – Vol. 45. – P. 290–298.

7. Семежев, Н. Моделирование диодных детекторов и их линеаризация / Н. Семежев, А. А. Львов, А. А. Солопекина // Компьютерные науки и информационные технологии : материалы Междунар. науч. конф. – Саратов : Издат. центр «Наука», 2016. – С. 364–368.

8. Lira, I. Bayesian Assessment of Uncertainty in Metrology: a Tutorial / I. Lira, D. Grientschnig // Metrologia. – 2010. – Vol. 47. – P. R1–R14.

9. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, Evaluation of Measurement Data—Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement'– Propagation of Distributions Using a Monte Carlo Method, Joint Committee for Guides in Metrology. – JCGM 101, 2008 – URL: http://www.bipm.org/utils/common /documents/jcgm/ JCGM_101_2008_E.pdf> (дата обращения: 29.08.25).

10. ГОСТ Р 54500.1–2011/Руководство ИСО/МЭК 98-3:2008. Неопределенность измерения. Ч. 3. Руководство по выражению неопределенности измерения. – Москва : Стандартинформ, 2011. – 57 с.

11. ГОСТ Р 34100.1–2017/ ISO/IEC Guide 98-1:2009. Неопределенность измерения. Часть 1. Введение в руководства по выражению неопределенности измерения. – Москва : Стандартинформ, 2018. – 22 с.

12. Солопекина, А. А. Расчет неопределенностей измерения характеристик многозондовой измерительной линии / А. А. Солопекина, А. А. Львов, Н. Семежев // Компьютерные науки и информационные технологии : материалы Междунар. науч. конф. – Саратов : Издат. центр «Наука», 2016. – С. 396–400.

13. Солопекина, А. А. Применение метода неопределенности для анализа погрешностей многополюсного рефлектометра / А. А. Солопекина, А. А. Львов, Н. Семежев, Н. С. Вагарина // Надежность и качество : сб. тр. Междунар. симп.: в 2 т. – Пенза : ПГУ, 2017. – Т. 2. – С. 136–139.

14. Львов, А. А. Анализ модели многозондовой измерительной линии и расчет неопределенностей измерения с ее помощью / А. А. Львов, Н. Семежев, А. А. Солопекина, О. М. Глухова // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика, 2019. – № 4. – С. 141–151. – DOI 10.24143/2072-9502-2019-4-141-151.

15. Solopekina, A. A. Calculation of Measurement Uncertainties of Multi-Port Transmission Line Reflectometer / A. A. Solopekina, A. A. Lvov, N. Semezhev // Proc. of the Int. Conf. on Actual Problems of Electron Devices Engineering. – Saratov : IEEE, 2014. – P. 356–362.

16. Solopekina, A. A. Application of the Uncertainty Method for Analysis of Multi-Port Correlator Accuracy / A. A. Solopekina, N. Semezhev, V. V. Komarov et al. // Proc. of the 2017 IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conf. – St. Petersburg : IEEE, 2016. – P. 505–509.

17. Kass, R. E. The Selection of Prior Distributions by Formal Rules / R. E. Kass, L. Wasserman // J. Am. Stat. Assoc. – 1996. – Iss. 91. – P. 1343–1370.

18. Berger, J. O. The Formal Definition of Reference Priors / J. O. Berger, J. M. Bernardo, D. Sun // Annals of Statistics. – 2009. – Vol. 37. – P. 905–938.

19. Yang, R. A Catalog of Noninformative Priors, ISDS Discussion paper 97-42 / R. Yang, J. Berger // Parexel Int. and Purdue Univ. – 1998. – URL: http://www.stats.org.uk/priors/noninformative/ YangBerger1998.pdf (дата обращения: 29.08.2025).

20. Box, G. E. P. Bayesian Inference in Statistical Analysis / G. E. P. Box, G. C. Tiao. – Addison Wesley Reading Mass, 1992. – 588 p.

21. Callegaro, L. Why Always Seek the Expected Value? A Discussion Relating to the Lp Norm / L. Callegaro, F. Pennecchi // Metrologia. – 2007. – Vol. 44. – P. L68–L70.

22. Gonin, R. Nonlinear Lp-norm Estimation / R. Gonin, A. H. Money. –New York and Basel : Marcel Dekker, Inc., 1989. – 316 p.

23. Kusnandar, N. Evaluation of Measurement Uncertainty Using the Monte Carlo Method in Steady-State Power Measurement of Household Refrigerator Based on IEC 62552:2015 / N. Kusnandar, Q. Lailiyah, I. Kasiyanto et al. // Evergreen Joint J. of Novel Carbon Resource Sciences & Green Asia Strategy. – 2024. – Vol. 11, iss. 03. – P. 2468–2478.

24. Машкин, В. А. Высокоточные измерения переменного тока, основанные на применении метода Монте-Карло / В. А. Машкин, А. А. Львов // Вестник Саратовского государственного технического университета. – 2009. № 2 (43), вып. 4. – С. 41–45.

25. Wubbeler, G. A Two-Stage Procedure for Determining the Number of Trials in the Application of a Monte Carlo Method for Uncertainty Evaluation / G. Wubbeler, P.M. Harris, M.G. Cox, C. Elster // Metrologia. – 2012. – Vol. 47. – P. 317– 324.

26. Линник, Ю. В. Метод наименьших квадратов и основы теории обработки наблюдений / Ю. В. Линник. – Москва : ГИФМЛ, 1958. – 334 с.

27. Вучков, И. Н. Прикладной линейный регрессионный анализ / И. Н. Вучков, Л. Н. Бояджиева, Е. Б. Солаков. – Москва : Финансы и статистика, 1987. – 239 с.

28. Мусатов, М. В. Анализ моделей метода наименьших квадратов и методов получения оценок / М. В. Мусатов, А. А. Львов // Вестник Саратовского государственного технического университета. – 2009. – № 2 (43), вып. 4. – С. 137–140.

29. Львов, А. А. Основы статистической обработки измерительной информации в задачах автоматического управления : учебное пособие для студентов высших учебных заведений, обучающихся по специальности 210100 – «Управление и информатика в технических системах» / А. А. Львов. – Саратов : СГТУ, 2005. – 84 с.

30. Van Huffel, S. Total Least Squares and Errors-in Variables Modeling: Analysis, Algorithms and Applications / S. Van Huffel, P. Lemmerling. – Kluwer Academic Publishers, Dordrecht, 2002. – 398 p.

31. Kariya, T. Generalized Least Squares / T. Kariya, H. Kurata. – Chichester : John Wiley & Sons, 2004. – 307 p.

32. Wolberg, J. Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments / J. Wolberg. – Berlin, Springer, 2005. 250 p.

33. Bernardo, J. M. Bayesian Methodology in Statistics / J. M. Bernardo // Comprehensive Chemometrics. – 2009. – Vol. 1. – P. 213–245.

34. Schantz, M. CCQM-K25: Determination of PCB Congeners in Sediment / M. Schantz, S. Wise // Metrologia. – 2004. – Vol. 41. – P. 08001 (Technical Supplement).

35. Рекомендация. Государственная система обеспечения единства измерений: Порядок реализации метрологическими институтами Росстандарта соглашения о взаимном признании национальных эталонов и сертификатов калибровки и измерений, выдаваемых национальными метрологическими институтами. МИ 3292-2010. – Москва : Стандратинформ, 2010. – 42 с.

36. Toman, B. Laboratory Effects Models for Interlaboratory Comparisons / B. Toman // Springer Nature. – 2009. – Iss. 14. – P. 553–563.

37. Абакумов, А. В. Исследование методов оценивания стандартного отклонения последовательности при контроле качества изделий / А. В. Абакумов, А. А. Львов, Е. Н. Скрипаль, Ю. А. Ульянина // Надежность и качество : сб. тр. Междунар. симп. : в 2 т. – Пенза: ПГУ, 2018. – Т. 2. – С. 92–96.

38. Comite International des Poids et Mesures (CIPM), Mutual Recognition of National Measurement Standards and of Calibration and Measurement Certificates issued by National Metrology Institutes, 1999, Technical report. – URL: http://www.bipm.org/pdf/ mra.pdf> (дата обращения: 29.08.25).

39. ГОСТ Р 34100.1-2017/ ISO/IEC Guide 98-1:2009. Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения (ISO/IEC Guide 98-3:2008, IDT). – Москва : Стандартинформ, 2017. – 105 с.

40. Laedermann, J. P. Measurement of Radioactive Samples: Application of the Bayesian Statistical Decision Theory / J. P. Laedermann, J. F. Valley, F. C. Bochud // Metrologia. – 2005. – Vol. 42. – P. 442–448.

References

1. *BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement. GUM 1995 with Minor Corrections, Joint Committee for Guides in Metrology*. JCGM 100, 2008. Available at: http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_ 2008_E.pdf.

2. *BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM), Joint Committee for Guides in Metrology*. JCGM 200, 2008, Available at: http://www.bipm.org/utils/common/documents/jcgm/ JCGM_200_2008.pdf.

3. *International Vocabulary of Metrology. Basic and General Concepts and Associated Terms.* Saint Petersburg, NPPO "Professional", 2010. 81 p. (In Russ.).

4. Balaban, O. M., Balaban, V. V., Lvov, A. A., Svetlov, M. S. Length measurements in Ancient Greece: Uncertainty of standards during the Golden Age of the Olympic Games. *Problems of Management, Processing, and Transmission of Information : Proceedings of the VI International Scientific Conference.* Saratov, LLC SOP "Lodi", 2019. pp. 611–620 (In Russ.).

5. Sharov, P. A., Umnova, E. G., Vagarina, N. S., Lvov, A. A., Svetlov, M. S. An alternative approach to expressing measurement uncertainty in experiments. *Mathematical Modeling, Computer and Natural Science Experiments*, 2021, no. 3. Available at: mathmod.esrae.ru/35-130 (accessed 27.08.2025). DOI 10.24412/2541-9269-2021-3-10-26 (In Russ.).

6. Willink, R. Estimation and Uncertainty in Fitting Straight Lines to Data: Different Techniques. *Metrologia*, 2008, vol. 45. pp. 290–298 (In Russ.).

7. Semezhev, N., Lvov, A. A., Solopekina, A. A. Modeling of diode detectors and their linearization. *Computer Science and Information Technologies : materials of the International Scientific Conference.* Saratov, Publishing Center "Nauka", 2016. pp. 364–368 (In Russ.).

8. Lira, I., Grientschnig, D. Bayesian Assessment of Uncertainty in Metrology: a tutorial. *Metrologia*, 2010, vol. 47, pp. R1–R14.

9. *BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, Evaluation of Measurement Data – Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement' – Propagation of Distributions Using a Monte Carlo Method, Joint Committee for Guides in Metrology.* JCGM: 101, 2008. Available at: http://www.bipm.org/utils/common/documents/jcgm/ JCGM_101_2008_E.pdf.

10. *GOST R 54500.1-2011 / ISO/IEC Guide 98-3:2008. Measurement Uncertainty. Part 3. Guide to Expression of Uncertainty in Measurement.* Moscow, Standartinform Publ., 2011. 57 p. (In Russ.).

11. *GOST R 34100.1-2017 / ISO/IEC Guide 98-1:2009. Measurement Uncertainty. Part 1. Introduction to the Guide to Expression of Uncertainty in Measurement.* Moscow, Standartinform, 2018. 22 p. (In Russ.).

12. Solopekina, A. A., Lvov, A. A. Semezhev, N. Calculation of measurement uncertainties of characteristics of a multi-probe measurement line. *Computer Science and Information Technologies : materials of the International Scientific Conference*. Saratov, Publishing Center "Nauka", 2016, pp. 396–400 (In Russ.).

13. Solopekina, A. A., Lvov, A. A., Semezhev, N., Vagarina, N. S. Application of the uncertainty method for analyzing errors of a multi-pole reflectometer. *Reliability and Quality : Proceedings of the International Symposium.* Penza, PSU, 2017, vol. 2, pp. 136–139 (In Russ.).

14. Lvov, A. A., Semezhev, N., Solopekina, A. A., Glukhova, O. M. Analysis of the multi-probe measurement line model and calculation of measurement uncertainties using it. *Bulletin of Astrakhan State Technical University. Series: Management, Computing and Informatics*, 2019, no. 4. pp. 141–151. DOI 10.24143/2072-9502-2019-4-141-151 (In Russ.).

15. Solopekina, A. A., Lvov, A. A., Semezhev, N. Calculation of Measurement Uncertainties of Multi-Port Transmission Line Reflectometer. *Proc. of the Int. Conf. on Actual Problems of Electron Devices Engineering*. Saratov, IEEE, 2014, pp. 356–362.

16. Solopekina, A. A., Semezhev, N., Komarov, V. V. et al. Application of the Uncertainty Method for Analysis of Multi-Port Correlator Accuracy. *Proc. of the 2017 IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conf.* St. Petersburg, IEEE, 2016, pp. 505–509.

17. Kass, R. E., Wasserman, L. The Selection of Prior Distributions by Formal Rules. *J. Am. Stat. Assoc.*, 1996, iss. 91, pp. 1343–1370.

18. Berger, J. O., Bernardo J. M., D. Sun The Formal Definition of Reference Priors. *Annals of Statistics*, 2009. Vol. 37, pp. 905–938.

19. Yang, R., Berger, J. A Catalog of Noninformative Priors, ISDS Discussion paper 97-42. *Parexel Int. and Purdue Univ.*, 1998. Available at: http://www.stats.org.uk/priors/noninformative/YangBerger1998.pdf (accessed 29.08.2025).

20. Box, G. E. P., Tiao, G. C. *Bayesian Inference in Statistical Analysis.* Addison Wesley Reading Mass, 1992. 588 p.

21. Callegaro, L., Pennecchi, F. Why Always Seek the Expected Value? A Discussion Relating to the Lp Norm. *Metrologia*, 2007, vol. 44, pp. L68–L70.

22. Gonin, R., Money, A. H. *Nonlinear Lp-norm Estimation.* New York and Basel, Marcel Dekker, Inc., 1989. 316 p.

23. Kusnandar, N., Lailiyah, Q., Kasiyanto, I. et al. Evaluation of Measurement Uncertainty Using the Monte Carlo Method in Steady-State Power Measurement of Household Refrigerator Based on IEC 62552:2015. *Evergreen Joint J. of Novel Carbon Resource Sciences & Green Asia Strategy*, 2024, vol. 11, iss. 03, pp. 2468–2478.

24. Mashkin, V. A., Lvov, A. A. High-precision measurements of alternating current based on the Monte Carlo method. *Bulletin of Saratov State Technical University*, 2009, no. 2 (43), issue 4, pp. 41–45 (In Russ.).

25. Wubbeler, G., Harris, P. M., Cox, M. G., Elster, C. A Two-Stage Procedure for Determining the Number of Trials in the Application of a Monte Carlo Method for Uncertainty Evaluation. *Metrologia*, 2012, vol. 47, pp. 317–324.

26. Linnik, Y. V. *The Least Squares Method and Fundamentals of Observation Processing.* Moscow, GIFML, 1958. 334 p. (In Russ.).

27. Vuchkov, I. N., Boyadzhieva, L. N., Solakov, E. B. *Applied Linear Regression Analysis*. Moscow, Finance and Statistics, 1987. 239 p. (In Russ.).

28. Musatov, M. V., Lvov, A. A. Analysis of the Least Squares Method models and estimation methods. *Bulletin of Saratov State Technical University*, 2009, no. 2 (43), issue 4, pp. 137–140 (In Russ.).

29. Lvov, A. A. *Fundamentals of Statistical Processing of Measurement Information in Automatic Control Tasks : textbook for university students majoring in 210100 – "Management and Informatics in Technical Systems".* Saratov, SSTU, 2005. 84 p. (In Russ.).

30. Van Huffel, S., Lemmerling, P. *Total Least Squares and Errors-in Variables Modeling: Analysis, Algorithms and Applications.* Dordrecht, Kluwer Academic Publishers, 2002. 398 p.

31. Kariya, T., Kurata, H. Generalized Least Squares. Chichester, John Wiley & Sons, 2004. 307 p.

32. Wolberg, J. *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments.* Berlin, Springer, 2005. 250 p.

33. Bernardo, J. M. Bayesian Methodology in Statistics. *Comprehensive Chemometrics*, 2009, vol. 1, pp. 213–245.

34. Schantz, M., Wise, S. CCQM-K25: Determination of PCB Congeners in Sediment. *Metrologia*, 2004, vol. 41, p. 08001 (Technical Supplement).

35. *Recommendation. State System for Ensuring the Uniformity of Measurements: Procedure for Implementation by Metrological Institutes of Rostransstandard of the Mutual Recognition of National Standards and Calibration and Measurement Certificates Issued by National Metrological Institutes. MI 3292-2010.* Moscow, Standartinform Publ., 2010. 42 p. (In Russ.).

36. Toman, B. Laboratory Effects Models for Interlaboratory Comparisons. *Springer Nature*, 2009, iss. 14, pp. 553–563.

37. Abakumov, A. V., Lvov, A. A., Skripal, E. N., Ulyanina, Y. A. Study of Methods for Estimating the Standard Deviation of a Sequence in Quality Control of Products. *Reliability and Quality : Collection of Proceedings of the International Symposium : in 2 vol.* Penza, PGSU, 2018, vol. 2, pp. 92–96 (In Russ.).

38. *Comite International des Poids et Mesures (CIPM), Mutual Recognition of National Measurement Standards and of Calibration and Measurement Certificates issued by National Metrology Institutes, 1999, Technical report*. Available at: http://www.bipm.org/pdf/ mra.pdf> (дата обращения: 29.08.2025).

39. *GOST R 34100.1-2017/ ISO/IEC Guide 98-1:2009. Uncertainty of measurement. Part 3. Guide to the expression of uncertainty in measurement (ISO/IEC Guide 98-3:2008, IDT).* Moscow, Standartinform Publ., 2017. 105 p. (In Russ.).

40. Laedermann, J. P., J. F. Valley, F. C. Bochud Measurement of Radioactive Samples: Application of the Bayesian Statistical Decision Theory. *Metrologia*, 2005, vol. 42, pp. 442–448.

Статья поступила в редакцию 30.09.2025; одобрена после рецензирования 15.12.2025; принята к публикации 15.12.2025.

The article was submitted 30.09.2025; approved after reviewing 15.12.2025; accepted for publication 15.12.2025.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

УДК 004.89

ВЫБОР ОПТИМАЛЬНОЙ ФУНКЦИИ ПРИНАДЛЕЖНОСТИ ДЛЯ НЕЧЕТКИХ МНОЖЕСТВ НА ОСНОВЕ ЭКСПЕРТНОЙ ОЦЕНКИ: СИСТЕМНЫЙ АНАЛИЗ, МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Соловьева Инна Александровна, Тамбовский государственный университет имени Г.Р. Державина, 392036, Российская Федерация, г. Тамбов, ул. Интернациональная, 33,
кандидат технических наук, доцент кафедры, ORCID: 0000-0002-1798-1859, e-mail: good.win32@yandex.ru
Соловьев Денис Сергеевич, Тамбовский государственный университет имени Г.Р. Державина, 392036, Российская Федерация, г. Тамбов, ул. Интернациональная, 33,
кандидат технических наук, доцент, ORCID: 0000-0001-6613-3218, e-mail: solovjevdenis @mail.ru
Горохова Светлана Александровна, Тамбовский государственный университет имени Г.Р. Державина, 392036, Российская Федерация, г. Тамбов, ул. Интернациональная, 33,
студент, ORCID: 0009-0009-8433-5281, e-mail: svetagorokhova13@gmail.com
Самохвалов Алексей Владимирович, Тамбовский государственный университет имени Г.Р. Державина, 392036, Российская Федерация, г. Тамбов, ул. Интернациональная, 33,
кандидат педагогических наук, доцент, ORCID: 0000-0002-3151-3250, e-mail: samohvalov @gmail.com

В статье рассматривается проблема выбора оптимальной функции принадлежности в задачах нечеткого моделирования. Для решения данной проблемы предлагается создание системы, которая автоматически определяет оптимальную параметрическую модель функции принадлежности на основе экспертных оценок, статистически оценивает ее адекватность и точность. Проведен системный анализ процесса построения функции принадлежности с использованием стандарта IDEF0. Представлены математические модели функций принадлежности, а также методы их оптимизации и статистической оценки. Описана программная реализация системы на языке Python, обеспечивающая автоматический подбор параметров, визуализацию и анализ результатов. Результаты вычислительных экспериментов подтверждают эффективность системы, демонстрируя, что выбор оптимальной функции принадлежности зависит от характера экспертных оценок.

Ключевые слова: оптимальный выбор, функция принадлежности, нечеткое множество, экспертная оценка, системный анализ, математическое моделирование, программная реализация

SELECTION OF THE OPTIMAL MEMBERSHIP FUNCTION FOR FUZZY SETS BASED ON EXPERT ASSESSMENT: SYSTEM ANALYSIS, MATHEMATICAL MODELING AND SOFTWARE IMPLEMENTATION

Solovjeva Inna A., Derzhavin Tambov State University, 33 Internatsionalnaya St., Tambov, 392036, Russian Federation,
Cand. Sci. (Engineering), Associate Professor of the Department, ORCID: 0000-0002-1798-1859, e-mail: good.win32@yandex.ru
Solovjev Denis S., Derzhavin Tambov State University, 33 Internatsionalnaya St., Tambov, 392036, Russian Federation,
Cand. Sci. (Engineering), Associate Professor, ORCID: 0000-0001-6613-3218, e-mail: solovjevdenis @mail.ru
Gorokhova Svetlana A., Derzhavin Tambov State University, 33 Internatsionalnaya St., Tambov, 392036, Russian Federation,
student, ORCID: 0009-0009-8433-5281, e-mail: svetagorokhova13@gmail.com
Samokhvalov Alexey V., Derzhavin Tambov State University, 33 Internatsionalnaya St., Tambov, 392036, Russian Federation,
Cand. Sci. (Pedagogics), Associate Professor, ORCID: 0000-0002-3151-3250, e-mail: samohvalov @gmail.com

The article addresses the problem of selecting an optimal membership function in fuzzy modeling tasks. To solve this problem, the development of a system is proposed that automatically determines the optimal parametric model of a membership function based on expert assessments, statistically evaluates its adequacy and accuracy. A systematic analysis of the membership function construction process is conducted using the IDEF0 methodology. Mathematical models of membership functions are presented, along with methods for their optimization and statistical evaluation.

The software implementation of the system in Python is described, ensuring automatic parameter selection, visualization, and result analysis. The results of computational experiments confirm the system's efficiency, demonstrating that the choice of the optimal membership function depends on the nature of expert assessments.

Keywords: optimal selection; membership function; fuzzy set; expert assessment; system analysis; mathematical modeling; software implementation

Graphical annotation (Графическая аннотация)

	1	2	3	4	5	6	7	8
1	Эксперты							
2	x	10	20	30	40	50	60	70
3	Эксперт 1	0	0	0	0	0	0	1
4	Эксперт 2	0	0	0	0	0	1	1
5	Эксперт 3	0	1	0	1	1	1	1
6	Эксперт 4	0	1	1	1	1	1	1
7	Эксперт 5	0	1	1	1	1	1	1
8	Эксперт 6	1	1	1	1	1	1	1
9	y	0,16667	0,66667	0,5	0,66667	0,66667	0,83333	1

ВВЕДЕНИЕ

В настоящее время, когда во многих областях науки и техники приходится работать с неполными или неточными данными, все больше внимания уделяется инструментам, способным учитывать неопределенность. Одним из таких инструментов является теория нечетких множеств, которая позволяет описывать реальность более гибко, чем традиционные бинарные методы [1]. Она дает возможность работать не только с «да» или «нет», но и с промежуточными значениями, что особенно важно в сложных системах, где границы между состояниями часто размыты [2]. В основе теории нечетких множеств лежит ключевое понятие – функция принадлежности, посредством которой можно численно описать, насколько объект соотносится с заданным множеством. В работе [3] продемонстрированы результаты прогнозирования валютного курса доллара США к рублю на основе аппарата нечетких множеств с колоколообразной (П-образной) функцией принадлежности. Методика организации тестирования и оценки результатов с применением теории нечетких множеств, использующей трапецеидальные функции принадлежности для формализации вербальных экспертных данных и учета времени ответов испытуемых, рассматривается в работе [4]. В работе [5] предложено применение функций принадлежности, основанных на нормальном распределении Гаусса, для поддержки принятия решений в условиях неопределенности, на примере определения приоритетного числа возможности улучшения в испытательной лаборатории. Использование нейронечеткой модели с треугольными функциями принадлежности для формирования базы знаний нечетко-продукционного типа в задаче подбора геолого-технических мероприятий на нефтяном месторождении описывается в работе [6]. Работа [7] посвящена применению *S*- и *Z*-образных функций принадлежности в нечеткой логике для построения математической модели оценки качества пушно-мехового полуфабриката, реализованной в среде Matlab для последующей интеграции в систему автоматизированной сортировки. Приведенные исследования демонстрируют широкий спектр применения отличающихся моделей функций принадлежности для решения задач в различных предметных областях. Однако выбор конкретного вида функции принадлежности часто основывается на субъективных предпочтениях экспертов или априорных предположениях о характере данных [8], что может привести к снижению точности и адекватности модели. Для минимизации субъективности и обеспечения статистической обоснованности выбора функции принадлежности необходима разработка системы, способной автоматически определять оптимальную модель на основе анализа характера экспертных данных. Такой подход позволит учитывать специфику распределения оценок, их динамику и разброс, обеспечивая минимальную погрешность аппроксимации и максимальную объяснительную способность модели функции принадлежности.

Целью работы является создание системы для построения функций принадлежности, которая на основе экспертных оценок выбирает оптимальную параметрическую модель, статистически оценивает ее адекватность и точность, а также визуализирует результаты для дальнейшего использования в задачах нечеткого моделирования.

СИСТЕМНЫЙ АНАЛИЗ ПРОЦЕССА ПОСТРОЕНИЯ ФУНКЦИИ ПРИНАДЛЕЖНОСТИ

Проведем системный анализ процесса построения функции принадлежности с применением стандарта IDEF0 [9]. Функциональная модель системы представлена на рисунке 1.

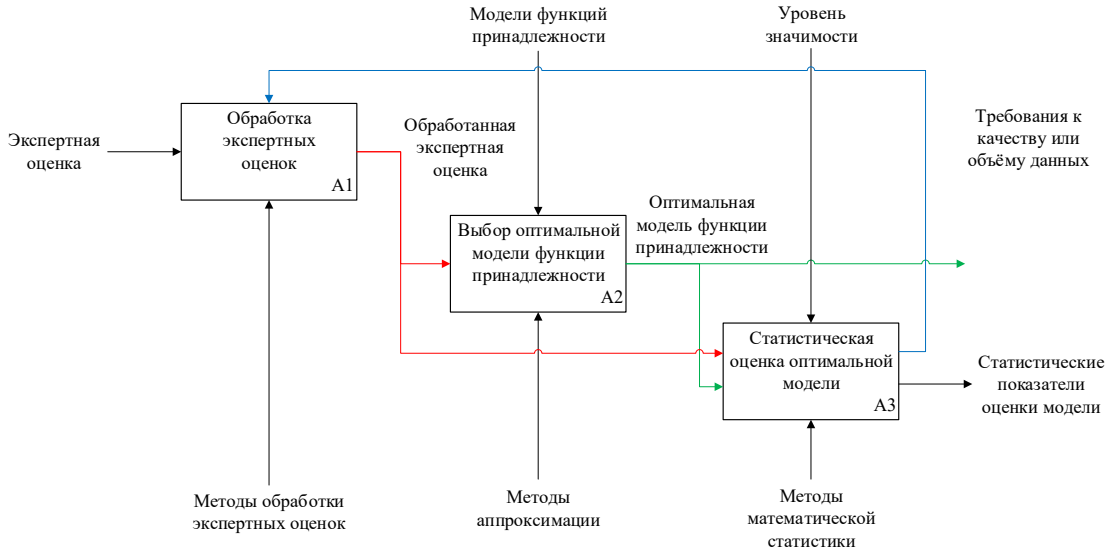


Рисунок 1 – Функциональная модель системы построения функции принадлежности в стандарте IDEF0

Модель системы, представленная в нотации IDEF0, включает три функциональных блока, связанных потоками данных и ограничениями.

Блок A1 «Обработка экспертных оценок» принимает на вход экспертные оценки и преобразует их с использованием специализированных методов обработки. Результатом является обработанная экспертная оценка, которая передается в блоки A2 «Выбор оптимальной модели функции принадлежности» и A3 «Статистическая оценка оптимальной модели». Работа блока A1 регламентируется требованиями к качеству и объему данных, формируемыми на выходе блока A3.

Блок A2 на основе обработанных экспертных оценок определяет оптимальную модель функции принадлежности, применяя методы аппроксимации. Выбор модели осуществляется из заданного множества допустимых вариантов функций принадлежности. Полученная оптимальная модель поступает на вход блока A3 для дальнейшего анализа.

Блок A3 выполняет статистическую оценку оптимальной модели, используя методы математической статистики. Входными данными являются обработанные экспертные оценки (от A1) и оптимальная модель функции принадлежности (от A2). Выходные данные включают статистические показатели оценки модели, а также требования к качеству или объему исходных данных, которые служат ограничением для блока A1. Корректировка требований направлена на достижение заданного уровня статистической значимости модели.

Таким образом, система реализует последовательный процесс обработки экспертных оценок, выбора оптимальной модели и ее статистической верификации, обеспечивая замкнутую обратную связь для контроля качества данных.

Рассмотрим более подробно ограничения (математические модели функций принадлежности) и механизмы (математические методы построения и статистической оценки функций принадлежности) системы.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ФУНКЦИЙ ПРИНАДЛЕЖНОСТИ

Существует множество типов функций принадлежности, каждая из которых позволяет моделировать различные лингвистические или логистические категории с определенной степенью размытости [10]. Так, типовыми параметрическими моделями функций принадлежности являются: *S*-образная, *Z*-образная, П-образная, треугольная, трапецеидальная, гауссова.

Аналитическое представление для *S*-образной функции принадлежности задается выражением:

$$f_1\left(x,\theta_1=\{a,b\}\right)=\frac{1}{1+e^{-a(x-b)}}, \quad (1)$$

где a, b – параметры, принимающие любые действительные значения при условии, что $a > 0$.

Аналитическое представление для *Z*-образной функции принадлежности задается выражением:

$$f_2\left(x,\theta_2=\{a,b\}\right)=\frac{1}{1+e^{a(x-b)}}, \tag{2}$$

где a, b – параметры, принимающие любые действительные значения при условии, что $a > 0$.

Аналитическое представление для П-образной функции принадлежности задается выражением:

$$f_3\left(x,\theta_3=\{a,b,c,d\}\right)=f_1\left(x,\theta_1=\{a,b\}\right)\cdot f_2\left(x,\theta_2=\{c,d\}\right), \tag{3}$$

где a, b, c, d – параметры, принимающие любые действительные значения при условии, что $a, c > 0$.

Аналитическое представление для треугольной функции принадлежности задается выражением:

$$f_4\left(x,\theta_4=\{a,b,c\}\right)=\begin{cases}0, x<a\\ \dfrac{x-a}{b-a}, a\le x<b\\ \dfrac{c-x}{c-b}, b\le x<c\\ 0, x\ge c\end{cases}, \tag{4}$$

где a, b, c – параметры, принимающие любые действительные значения при условии, что $a \le b \le c$.

Аналитическое представление для трапецеидальной функции принадлежности задается выражением:

$$f_5\left(x,\theta_5=\{a,b,c,d\}\right)=\begin{cases}0, x<a\\ \dfrac{x-a}{b-a}, a\le x<b\\ 1, b\le x<c\\ \dfrac{d-x}{d-c}, c\le x<d\\ 0, x\ge d\end{cases}, \tag{5}$$

где a, b, c, d – параметры, принимающие любые действительные значения при условии $a \le b \le c \le d$.

Аналитическое представление для гауссовой функции принадлежности задается выражением:

$$f_6\left(x,\theta_6=\{a,b\}\right)=e^{\frac{-(x-a)^2}{2b^2}}, \tag{6}$$

где a – параметр, имеющий смысл математического ожидания;

b – параметр, имеющий смысл среднеквадратического отклонения и принимающий значение $b > 0$.

Для построения функций принадлежности (1)–(6) с оптимальными параметрами применяются методы численной оптимизации и математической статистики.

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ПОСТРОЕНИЯ И СТАТИСТИЧЕСКОЙ ОЦЕНКИ ФУНКЦИЙ ПРИНАДЛЕЖНОСТИ

Дано множество экспертных оценок значений функции принадлежности:

$$\tilde{\mu}_{\cdot,\cdot,\cdot}, \tag{7}$$

на основе которого вычисляются усредненные значения:

$$\mu(x_i)=\frac{1}{M}\sum_{j=1}^{M}\tilde{\mu}_{\cdot,\cdot,\cdot}, \tag{8}$$

где x – значение оцениваемой переменной;

i – индекс оцениваемого значения ($i = 1,\ldots, N$);

j – индекс эксперта ($j = 1,\ldots, M$).

Имеется набор параметрических моделей $\mathcal{M} = \{f_k(x,\theta_k)\}$, ($k = 1 ,\ldots, K$), среди которых S- (1), Z- (2), П-образная (3), треугольная (4), трапецеидальная (5) и гауссова (6) функции принадлежности.

Требуется для каждой модели f_k с учетом (7), (8) найти параметры θ_k^*, минимизирующие среднеквадратическое отклонение:

$$\theta_k^*=\underset{\theta_k}{\operatorname{argmin}}\,SSE\left(\theta_k\right)=\underset{\theta_k}{\operatorname{argmin}}\sum_{i=1}^{N}\left(\mu\left(x_i\right)-f_k\left(x_i,\theta_k\right)\right)^2. \tag{9}$$

После расчета (9) необходимо выбрать модель f_{k*}, обеспечивающую наименьшую ошибку:

$$k^*=\underset{k}{\operatorname{argmin}}\,SSE\left(\theta_k^*\right)=\underset{k}{\operatorname{argmin}}\sum_{i=1}^{N}\left(\mu\left(x_i\right)-f_k\left(x_i,\theta_k^*\right)\right)^2. \tag{10}$$

Выбранная согласно (10) модель f_{k*} проверяется на адекватность с использованием F-критерия Фишера [11]:

$$F = \frac{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\tilde{\mu}_{\ldots}(x_i)\right)^2}{\frac{1}{N-\dim\left(\theta_{k^*}^*\right)}\sum_{i=1}^{N}\left(\mu(x_i)-f_k\left(x_i,\theta_{k^*}^*\right)\right)^2}. \quad (11)$$

Если расчетное значение (11) превышает табличное значение F-критерия Фишера $F_{\text{табл}}(\alpha, \dim(\theta_{k^*}^*), N-\dim(\theta_{k^*}^*))$, то модель f_{k*} признается адекватной на уровне значимости α с долей объясненной дисперсии:

$$R^2 = 1 - \frac{SSE\left(\theta_{k^*}^*\right)}{\sum_{i=1}^{N}\left(\mu(x_i)-\frac{1}{N}\sum_{i=1}^{N}\mu(x_i)\right)^2}. \quad (12)$$

Таким образом, оптимальной выбранной моделью функции принадлежности признается модель с минимальной SSE (10), прошедшая F-тест (11) и имеющая R^2 (12) ~ 1.

Для найденных оптимальных параметров $\theta_{k^*}^*$ модели f_{k*} вычисляются доверительные интервалы с использованием t-критерия Стьюдента:

$$\theta_{k^*,l}^* \pm t_{1-\alpha/2,N-\dim\left(\theta_{k^*}^*\right)} SE\left(\theta_{k^*,l}^*\right), \quad (13)$$

где $t_{1-\alpha/2,N-\dim\left(\theta_{k^*}^*\right)}$ – табличное значение t-распределения Стьюдента для заданного уровня значимости α и числа степеней свободы $N-\dim(\theta_{k^*}^*)$;

$SE\left(\theta_{k^*,l}^*\right)$ – стандартная ошибка оценки параметра $\theta_{k^*,l}^*$, вычисляемая на основе ковариационной матрицы оценок; l – индекс параметра ($l = 1,\ldots, \dim(\theta_{k^*}^*)$).

Рассмотрим более подробно программную реализацию системы.

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ ПОСТРОЕНИЯ ФУНКЦИИ ПРИНАДЛЕЖНОСТИ

Для реализации системы построения функций принадлежности выбран язык программирования Python. Данный выбор обусловлен поддержкой множества библиотек для обработки данных, реализации методов численной оптимизации и статистических расчетов, графической визуализации данных и разработки пользовательских интерфейсов.

Представленный фрагмент кода реализует загрузку исходных данных из файла формата Excel с их последующей обработкой при помощи библиотеки Pandas. В случае отсутствия файла или некорректного формата данных выполнение кода завершается и выводится соответствующее сообщение об ошибке:

```
try:
    data = pd.read_excel('train.xlsx', header=None)
    x_data = data.iloc[1, 1:].values.astype(float)
    y_data = data.iloc[-1, 1:].values.astype(float)
except Exception as e:
    print(f"Ошибка при чтении данных из Excel файла: {e}")
    x_data = np.array([])
    y_data = np.array([])
    print("Данные не были загружены. Программа будет завершена.")
    exit()
if x_data.size == 0 or y_data.size == 0:
    print("Нет доступных данных для аппроксимации. Программа будет завершена.")
    exit()
```

Функции принадлежности задаются списком в виде лямбда-выражений, сопровождаемых метаданными для оптимизации:

```
functions = [
    {
        'func': lambda x, a, b: 1 / (1 + np.exp(-a * (x - b))),
        'name': 'S-образная функция',
        'p0': [1, np.median(x_data)],
        'bounds': ([0, -np.inf], [np.inf, np.inf])
    },
```

```
…
    {
        'func': lambda x, a, b: np.exp(-(x - a)**2 / (2 * b**2)),
        'name': 'Гауссова функция',
        'p0': [np.median(x_data), (x_data.max() - x_data.min())/4],
        'bounds': ([-np.inf, 0], [np.inf, np.inf])
    },
]
```

Элементы списка представлены словарями с параметрами: 1) func – реализация функции, передаваемой алгоритму оптимизации; 2) name – название функции для интерпретации результатов в отчетах и графиках; 3) p0 – начальные значения параметров, определяемые на основе статистических характеристик исходных данных (минимальное, максимальное, среднее, среднеквадратическое); 4) bounds – границы допустимых значений параметров, исключающие физически некорректные решения.

Процедура построения функций принадлежности реализована в функции fit_data, выполнение которой начинается с определения списка функций, которые пользователь выбрал для аппроксимации:

```
def fit_data():
    selected_functions = [f for f, var in zip(functions, check_vars) if var.get()]
    results = []
```

Далее вызывается функция curve_fit из библиотеки SciPy, которая реализует подбор оптимальных параметров по алгоритму Левенберга – Марквардта:

```
    for f in selected_functions:
        params, covariance = curve_fit(f['func'], x_data, y_data, p0=f['p0'], bounds=f['bounds'])
```

Представленный фрагмент кода выполняет статистический анализ параметрической модели функций принадлежности и сохранение полученных результатов в структурированном виде:

```
        y_fit = f['func'](x_data, *params)
        residuals = y_data - y_fit
        ss_res = np.sum(residuals ** 2)
        ss_tot = np.sum((y_data - np.mean(y_data)) ** 2)
        r_squared = 1 - (ss_res / ss_tot)
        n = len(y_data)
        p = len(params)
        perr = np.sqrt(np.diag(covariance))
        t_values = params / perr
        p_values = [2 * (1 - stats.t.cdf(np.abs(t_val), n - p)) for t_val in t_values]
        df_model = p
        df_resid = n - p
        ss_reg = ss_tot - ss_res
        ms_reg = ss_reg / df_model if df_model != 0 else 0
        mse = ss_res / df_resid if df_resid != 0 else 0
        if ss_reg <= 0 or mse == 0:
            f_stat_model = 0
            p_value_model = 1.0
        else:
            f_stat_model = ms_reg / mse
            p_value_model = 1 - stats.f.cdf(f_stat_model, df_model, df_resid)
        results.append({
            'name': f['name'],
            'params': params,
            'param_errors': perr,
            't_values': t_values,
            'p_values': p_values,
            'y_fit': y_fit,
            'ss_res': ss_res,
            'r_squared': r_squared,
            'f_stat_model': f_stat_model,
            'p_value_model': p_value_model,
            'fit_time': (fit_end_time - fit_start_time).total_seconds(),
            'df_model': df_model,
            'df_resid': df_resid
        })
```

```
alpha = 0.05
if results:
    best_fit = min(results, key=lambda x: x['ss_res'])
    df_model_best = best_fit['df_model']
    df_resid_best = best_fit['df_resid']
    f_critical = stats.f.ppf(1 - alpha, df_model_best, df_resid_best)
    if best_fit['f_stat_model'] > f_critical:
        output_text.insert(tk.END, "Модель значима на уровне значимости 0,05.\n")
    else:
        output_text.insert(tk.END, "Модель не значима на уровне значимости 0,05.\n Рекомендация: добавьте больше точек и/или измените входные данные\n ")
```

Для графического отображения результатов используется библиотека Matplotlib. Исходные данные отображаются точками, аппроксимирующие кривые – гладкими линиями. Настройки осей и легенды обеспечивают наглядность сравнения моделей функций принадлежности:

```
x_smooth = np.linspace(min(x_data), max(x_data), 1000)
plt.figure(figsize=(10, 6))
plt.plot(x_data, y_data, 'o', label='Исходные данные')
for res in results:
    func = next(f['func'] for f in functions if f['name'] == res['name'])
    y_smooth = func(x_smooth, *res['params'])
    plt.plot(x_smooth, y_smooth, label=f"{res['name']}")
plt.xticks(range(int(min(x_data)), int(max(x_data)) + 1, 5))
plt.yticks(np.arange(0, 1.1, 0.1))
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left', borderaxespad=0)
plt.xlabel('x')
plt.ylabel('y')
plt.grid(True, linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

Интерфейс, реализованный на базе Tkinter, включает: 1) чекбоксы для выбора функций принадлежности; 2) кнопки запуска расчета и сохранения полученных результатов; 3) текстовое поле для вывода результатов аппроксимации и статистического анализа.

```
root = tk.Tk()
root.title("Выбор функций для аппроксимации")
check_vars = [tk.BooleanVar(value=True) for _ in functions]
for i, f in enumerate(functions):
    tk.Checkbutton(root, text=f['name'], variable=check_vars[i]).grid(row=i, column=0, sticky=tk.W)
tk.Button(root, text="Старт", command=fit_data).grid(row=len(functions), column=0, pady=10)
tk.Button(root, text="Сохранить результаты", command=save_results).grid(row=len(functions) + 1, column=0, pady=10)
output_text = tk.Text(root, height=20, width=80)
output_text.grid(row=0, column=1, rowspan=len(functions) + 2, padx=10, pady=10)
root.mainloop()
```

Таким образом, представленная программная реализация обеспечивает автоматический подбор и оценку функций принадлежности на основе обработки экспертных данных, включая статистический анализ значимости моделей и их визуализацию.

ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ И АНАЛИЗ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

В рамках данного исследования проведена серия вычислительных экспериментов, направленных на демонстрацию работоспособности разработанной системы построения функций принадлежности.

На рисунке 2 приведены результаты оценки шестью экспертами семи значений функций принадлежности, определенные на заданном диапазоне от 10 до 70, для трех различных случаев. Исходные данные экспериментов отражают типичные сценарии, встречающиеся в задачах нечеткого моделирования: от четко выраженных монотонных изменений до сложных нелинейных зависимостей.

	1	2	3	4	5	6	7
1	Эксперты						
2	x	10	20	30	40	50	60
3	Эксперт 1	0	0	0	0	0	0
4	Эксперт 2	0	0	0	0	0	1
5	Эксперт 3	0	1	0	1	1	1
6	Эксперт 4	0	1	1	1	1	1
7	Эксперт 5	0	1	1	1	1	1
8	Эксперт 6	1	1	1	1	1	1
9	y	0,16667	0,66667	0,5	0,66667	0,66667	0,83333

а)

	1	2	3	4	5	6	7
1	Эксперты						
2	x	10	20	30	40	50	60
3	Эксперт 1	1	0	0	0	0	0
4	Эксперт 2	1	1	1	0	0	0
5	Эксперт 3	1	1	1	0	1	0
6	Эксперт 4	1	1	1	1	1	0
7	Эксперт 5	1	1	1	1	1	0
8	Эксперт 6	1	1	1	1	1	1
9	y	1	0,83333	0,83333	0,5	0,66667	0,16667

б)

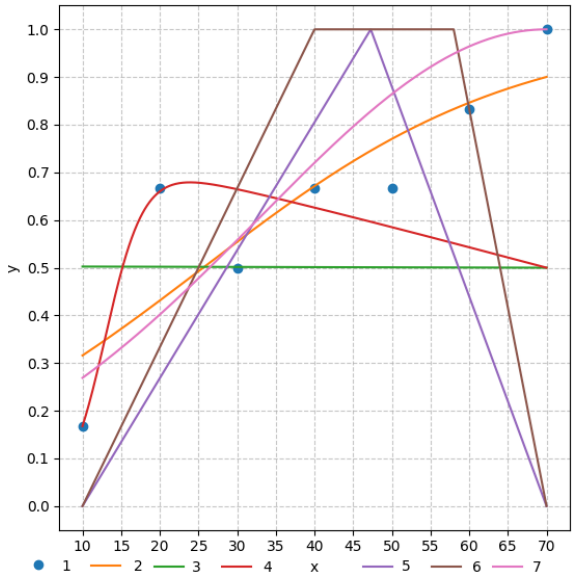
	1	2	3	4	5	6	7	8
1	Эксперты							
2	x	10	20	30	40	50	60	70
3	Эксперт 1	0	0	0	1	0	0	0
4	Эксперт 2	0	0	0	1	1	0	0
5	Эксперт 3	0	0	1	1	1	1	0
6	Эксперт 4	0	1	1	1	1	1	0
7	Эксперт 5	1	1	1	1	1	1	1
8	Эксперт 6	1	1	1	1	1	1	1
9	y	0,33333	0,5	0,66667	1	0,83333	0,66667	0,33333

в)

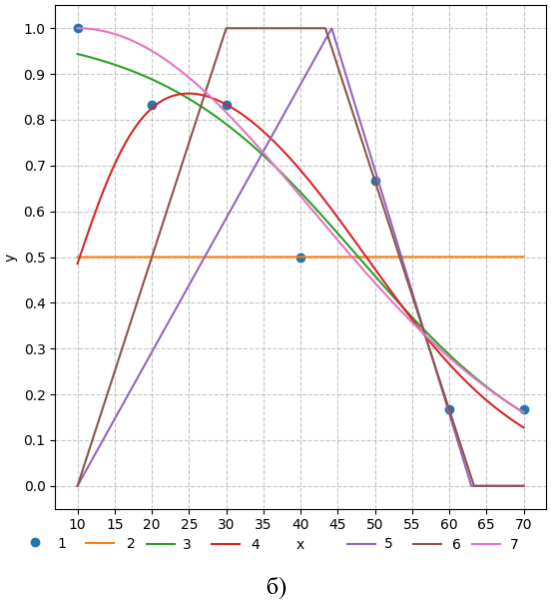
Рисунок 2 – Экспертные оценки значений функций принадлежности для экспериментов: а) 1; б) 2; в) 3

В эксперименте 1 экспертные оценки демонстрируют высокую степень проявления исследуемых признаков. В эксперименте 2 наблюдается слабая степень выраженности оцениваемых признаков. В эксперименте 3 присутствует динамика изменений: первоначально высокая степень проявления признаков сменяется их слабой выраженностью.

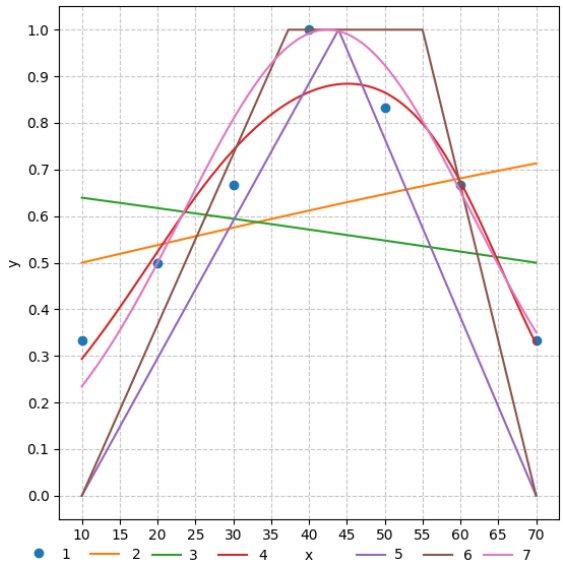
На рисунке 3 приведены сформированные в результате решения задачи аппроксимации функции принадлежности для трех различных случаев. Здесь приняты следующие обозначения: 1 – исходные данные; 2 – *S*-образная; 3 – *Z*-образная; 4 – П-образная; 5 – треугольная; 6 – трапецеидальная; 7 – гауссова функции. В каждом случае оптимальная функция принадлежности отличается.



а)



б)



в)

Рисунок 3 – Сформированные в результате решения задачи аппроксимации функции принадлежности для экспериментов: а) 1; б) 2; в) 3

В таблице приведены численные результаты выбора оптимальной функции принадлежности для экспертных оценок из рассматриваемых экспериментов.

Таблица – Численные результаты выбора оптимальной функции принадлежности

Параметры и показатели		№ эксперимента		
		1	2	3
$\theta^*_{k^*}$	*a*	25,6±5,9	47,8±3,8	19,0±2,6
	b	20,2±7,0	13,4±3,8	10,2±3,0
	c	–	–	65,2±2,0
	d	–	–	6,7±2,3
SSE		0,102	0,086	0,027
F		7,66	16,72	9,69
$F_{табл}$		5,79	5,79	9,12
R^2		0,754	0,870	0,928

В эксперименте 1 плавный переход от низких к высоким значениям принадлежности признаков наилучше описывается *S*-образной функцией. Убывание, характерное для слабой выраженности признаков в эксперименте 2, наилучшим образом моделирует *Z*-образная функция. В эксперименте 3 с пиком выраженности признака в среднем диапазоне значений наилучшее соответствие демонстрирует П-образная функция принадлежности. Расчетные значения *F*-критерия для всех результатов превышают свои табличные значения, что подтверждает адекватность параметрических моделей функций принадлежности на уровне статистической значимости α = 0,05. Коэффициент детерминации R^2 достигает наивысшего значения в эксперименте 3, что свидетельствует о высокой точности аппроксимации П-образной функцией принадлежности, тогда как в экспериментах 1 и 2 его значения указывают на меньшую, но все же значительную объяснительную способность *S*-образной и *Z*-образной функций. Данное обстоятельство подтверждается уменьшением ширины доверительного интервала для найденных оптимальных значений параметров выбранных функций принадлежности с уменьшением разброса экспертных оценок.

Треугольная, трапецеидальная и гауссова функции показали меньшую точность в данных экспериментах, что связано с их ограниченной гибкостью при описании переходов и сложных динамик.

ЗАКЛЮЧЕНИЕ

Проведенное исследование демонстрирует эффективность разработанной системы для построения функций принадлежности на основе экспертных оценок. Система обеспечивает автоматический подбор оптимальных параметрических моделей, их статистическую оценку и визуализацию, что позволяет минимизировать субъективность и повысить точность результатов. Результаты вычислительных экспериментов подтверждают необходимость учета специфики данных при выборе для них модели функции принадлежности. Разработанная система обладает практической значимостью для задач принятия решений в условиях неопределенности, классификации объектов с нечеткими

границами и моделирования сложных систем. Дальнейшее развитие системы может включать расширение библиотеки функций принадлежности и интеграцию методов машинного обучения для обработки больших объемов качественных данных, а также адаптацию алгоритмов для работы с многомерными нечеткими множествами.

Список источников

1. Zadeh, L. A. Fuzzy logic / L. A. Zadeh // Computer. – 1988. – Vol. 21 (4). – P. 83–93.

2. Bilgiç, T., Measurement of Membership Functions: Theoretical and Empirical Work / T. Bilgiç, I. B. Türkşen // The Handbooks of Fuzzy Sets Series. – 2000. – Vol 7. – P. 195–227.

3. Пилюгина, А. В. Опыт использования аппарата нечетких множеств в прогнозировании валютного курса / А. В. Пилюгина, А. А. Бойко // Прикаспийский журнал: управление и высокие технологии. – 2014. – № 3 (27). – С. 143–157.

4. Станишевская, А. В. Снижение уровня субъективной неопределенности при использовании тестирования для оценки уровня компетентности испытуемых / А. В. Станишевская, И. М. Ажмухамедов // Прикаспийский журнал: управление и высокие технологии. – 2019. – № 2 (46). – С. 153–162.

5. Аль-Бусаиди, С. С. С. К вопросу о поддержке процесса принятия решения об улучшении деятельности в испытательной лаборатории / С. С. С. Аль-Бусаиди, Ю. Н. Воякина, С. В. Пономарев // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 1 (53). – С. 27–45.

6. Катасева, Д. В. Нейронечеткая модель и программный комплекс формирования баз знаний для оценки состояния объектов / Д. В. Катасева // Прикаспийский журнал: управление и высокие технологии. – 2022. – № 1 (57). – С. 65–76.

7. Бодрякова, Л. Н. Применение нечеткой логики в технологических процессах изготовления швейных изделий / Л. Н. Бодрякова, И. С. Беляев, А. В. Фаддеенков [и др.] // Дизайн и технологии. – 2024. – № 102 (144). – С. 28–38.

8. Каид, В. А. А. Методы построения функций принадлежности нечетких множеств / В. А. А. Каид // Известия ЮФУ. Технические науки. – 2013. – № 2 (139). – С.144–153.

9. Методология функционального моделирования IDEF0. Руководящий документ РД IDEF0-2000. – Москва : Госстандарт России, 2000. – 75 с.

10. Dubois, D. Membership Functions / D. Dubois, H. Prade // Studies in Fuzziness and Soft Computing. – 2021. – Vol. 394. – P. 5–20.

11. Шириков, В. Ф. Математическая статистика / В. Ф. Шириков, С. М. Зарбалиев. – Москва : КолосС, 2009. – 480 с.

References

1. Zadeh, L. A. Fuzzy logic. *Computer*, 1988, vol. 21 (4), pp. 83–93.

2. Bilgiç, T., Türkşen, I. B. Measurement of Membership Functions: Theoretical and Empirical Work. *The Handbooks of Fuzzy Sets Series*, 2000, vol. 7, pp. 195–227.

3. Pilyugina, A. V., Boiko, A. A. Experience with using fuzzy sets in forecasting of currency exchange rate. *Caspian Journal: Control and High Technologies*, 2014, no. 3 (27), pp. 143–157 (In Russ.).

4. Stanishevskaya, A. V., Azhmukhamedov, I. M. The decrease of the subjective uncertainty level during the use of testing to estimate the level of competence of the tested. *Caspian Journal: Control and High Technologies*, 2019, no. 2 (46), pp. 153–162 (In Russ.).

5. Al-Busaidi, S. S., Voyakina, Yu. N., Ponomarev, S. V. To the question of supporting the decision-making process on performance improvement in the testing laboratory. *Caspian Journal: Control and High Technologies*, 2021, no. 1 (53), pp. 27–45 (In Russ.).

6. Kataseva, D. V. Neuro-fuzzy model and software complex for forming knowledge bases for objects state assessing. *Caspian Journal: Control and High Technologies*, 2022, no. 1 (57), pp. 65–76 (In Russ.).

7. Bodryakova, L. N., Belyaev, I. S., Faddeenkov, A. V., Tennikov, L. E., Ibadov, I. U. Application of fuzzy logic in technological processes of garment manufacturing. *Design and Technology*, 2024, no. 104 (144), pp. 28–38 (In Russ.).

8. Qaid, W. A. A. Methods construction membership function of fuzzy sets. *Izvestiya SFedU. Engineering Sciences*, 2013, no. 2 (139), pp. 144–153 (In Russ.).

9. *Methodology of functional modeling IDEF0. Guidance document RD IDEF0-2000.* Moscow, State Standard of Russia, 2000. 75 p. (In Russ.).

10. Dubois, D., Prade, H. Membership Functions. *Studies in Fuzziness and Soft Computing*, 2021, vol. 394, pp. 5–20.

11. Shirikov, V. F., Zarbaliyev, S. M. *Mathematical statistics.* Moscow, KolosS Publ., 2009. 480 p. (In Russ.).

Статья поступила в редакцию 10.07.2025; одобрена после рецензирования 25.07.2025; принята к публикации 05.08.2025.

The article was submitted 10.07.2025; approved after reviewing 25.07.2025; accepted for publication 05.08.2025.

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

УДК 004.056.5, 003.26

МОДИФИЦИРОВАННЫЙ МЕТОД ЧАСТОТНОГО КРИПТОАНАЛИЗА ШИФРА ВЕРТИКАЛЬНОЙ ПЕРЕСТАНОВКИ

Демина Раиса Юрьевна, Астраханский государственный технический университет, 414025, Российская Федерация, г. Астрахань, ул. Татищева, 16,

кандидат технических наук, доцент, ORCID: 0009-0009-1615-5641, e-mail: raisa.demina.91@mail.ru

Хайтул Анастасия Всеволодовна, Астраханский государственный университет им. В. Н. Татищева, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

магистрант, ORCID: 0000-0003-2112-8145, e-mail: khaaaytul@icloud.com

В статье рассматривается методика криптоанализа шифра вертикальной перестановки, основанная на частотном анализе триграмм. Предложен способ сокращения переборного пространства при взломе путтм исключения маловероятных комбинаций столбцов на основе частотных характеристик естественного языка. Метод показал существенное сокращение количества итераций при сохранении достоверности расшифровки. Работа ориентирована на повышение эффективности обучения студентов направлений, связанных с информационной безопасностью.

Ключевые слова: криптоанализ, частотный анализ, вертикальная перестановка, триграммы, шифр, информационная безопасность

A MODIFIED METHOD OF FREQUENCY CRYPTANALYSIS OF A VERTICAL PERMUTATION CIPHER

***Demina Raisa Yu.*,** Astrakhan State Technical University, 16 Tatishchev St., Astrakhan, 414025, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, ORCID: 0009-0009-1615-5641, e-mail: raisa.demina.91@mail.ru

Khaytul Anastasia V., Astrakhan Tatishchev State University, 20a Tatishchev St., 414056, Russian Federation,

master's student, ORCID: 0000-0003-2112-8145, e-mail: khaaaytul@icloud.com

The article discusses a technique for cryptanalysis of a vertical permutation cipher based on frequency analysis of trigrams. A method is proposed to reduce the bulkhead space during hacking by eliminating unlikely column combinations based on the frequency characteristics of the natural language. The method showed a significant reduction in the number of iterations while maintaining the reliability of the decryption. The work is focused on improving the effectiveness of students' education in areas related to information security.

Keywords: cryptanalysis, frequency analysis, vertical permutation, trigrams, cipher, information security

ВВЕДЕНИЕ

Подготовка квалифицированных специалистов в области информационной безопасности требует развития не только технических знаний и навыков, но и способности к системному и критическому мышлению. Существенным элементом профессиональной подготовки является умение рассматривать информационную систему как с точки зрения ее конструктивной защиты, так и с позиции потенциального нарушителя. Такой двусторонний подход позволяет будущим специалистам по защите информации своевременно выявлять уязвимости и оценивать степень устойчивости используемых методов криптографической защиты.

Дисциплина «Криптографические методы защиты информации» играет ключевую роль в формировании этих компетенций. В процессе обучения студенты знакомятся с широким спектром алгоритмов шифрования – от современных шифров до устаревших, имеющих в основном учебно-методическое значение. Особое внимание уделяется поиску уязвимостей этих алгоритмов, что способствует развитию у студентов навыков глубокого анализа, логического рассуждения и оценки потенциальной стойкости систем.

Учитывая высокую сложность современных симметричных криптоалгоритмов, их практический взлом невозможен без использования значительных вычислительных ресурсов. В связи с этим в образовательной практике целесообразно использовать алгоритмы шифрования докомпьютерной эпохи. Их структура и принципы функционирования позволяют наглядно продемонстрировать основные методы криптоанализа и научить студентов применять их в различных ситуациях.

Наиболее широко в учебных курсах рассматриваются шифры замены (Цезаря, Виженера, Хилла и др.), при этом они достаточно эффективно поддаются частотному анализу. Вместе с тем шифры перестановки, в том числе вертикальный шифр (шифр столбцовой перестановки), в образовательных материалах зачастую анализируются поверхностно. Традиционно предлагаемый метод их криптоанализа сводится к перебору возможных комбинаций столбцов, что мало способствует развитию аналитического мышления у обучающихся и требует значительных затрат времени при увеличении длины ключа.

В настоящей работе рассматривается применение частотного криптоанализа к шифру вертикальной перестановки. Предлагаемый подход позволяет существенно сократить пространство перебора за счет исключения комбинаций, содержащих маловероятные для естественного языка триграммы. Применение такого метода не только повышает эффективность криптоанализа, но и способствует формированию у студентов нестандартного мышления и способности адаптировать известные инструменты к новым задачам.

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ШИФРА ВЕРТИКАЛЬНОЙ ПЕРЕСТАНОВКИ

Шифры перестановки представляют собой один из наиболее ранних классов симметричных криптосистем, в которых сохраняется состав алфавита открытого текста, но изменяется порядок следования символов. В отличие от шифров замены, где символы заменяются на другие, перестановочные шифры обеспечивают конфиденциальность исключительно за счет реорганизации текста. Это делает их особенно удобными для использования в учебной практике, поскольку они наглядно демонстрируют зависимость криптостойкости от порядка элементов и длины ключа [1–2].

Одним из типичных представителей данного класса является шифр вертикальной перестановки (также известный как шифр столбцовой перестановки). Его принцип заключается в том, что открытый текст записывается по строкам в прямоугольную таблицу фиксированной ширины (определяемой длиной ключа), после чего столбцы таблицы переставляются согласно ключевому порядку. Итоговый шифртекст формируется путем последовательного считывания символов по столбцам, в новом порядке [3].

Формально процесс можно описать следующим образом. Пусть имеется открытый текст T длиной N символов и ключ перестановки K, задающий новую последовательность столбцов. Текст записывается в таблицу размером $r \times c$, где $c = |K|$, а $r = \lceil N/c \rceil$. Если N не кратно c, оставшиеся ячейки заполняются специальными символами-заполнителями (например, пробелами или другими знаками). Далее столбцы переставляются в соответствии с порядком, заданным ключом K, и шифртекст считывается по столбцам сверху вниз.

Например, при использовании ключа длиной 6 и шифруемом тексте длиной 42 символа шифровальная таблица будет иметь 7 строк и 6 столбцов. Если ключ определяет порядок столбцов как (3, 1, 4, 6, 2, 5), то третий столбец станет первым, первый – вторым и т. д.

Стойкость данного шифра определяется числом возможных перестановок столбцов. При длине ключа c общее число таких перестановок равно $c!$. Таким образом, уже при $c = 10$ возникает более 3,6 миллиона возможных вариантов (точнее, 10! = 3 628 800), что делает полный перебор крайне неэффективным даже для учебных целей [4].

Несмотря на относительную простоту алгоритма, анализ и взлом шифра вертикальной перестановки при помощи автоматизированных методов требует от студентов умения формализовывать задачу, учитывать вероятностные свойства языка и применять методы анализа структуры текста. Все это делает данный шифр удобной моделью для формирования практических навыков криптоанализа.

ТРАДИЦИОННЫЙ МЕТОД КРИПТОАНАЛИЗА ШИФРА ВЕРТИКАЛЬНОЙ ПЕРЕСТАНОВКИ

В учебной и методической литературе по криптографии классическим подходом к взлому шифра вертикальной перестановки считается метод полного или частичного перебора возможных вариантов упорядочивания столбцов шифровальной таблицы. Этот подход прост для реализации, но крайне ресурсоемок при увеличении длины ключа, поскольку требует перебора всех возможных перестановок столбцов, число которых при длине ключа n составляет $n!$ [5–6].

Процесс традиционного криптоанализа вертикального шифра обычно включает следующие шаги:

1. Определение длины шифртекста и вычисление возможных размерностей таблицы.
2. Выбор предполагаемой длины ключа, исходя из анализа структуры сообщения (в том числе, если известно, что ключ представляет собой слово).
3. Формирование таблицы с предполагаемой длиной ключа.
4. Перестановка столбцов таблицы во всех возможных порядках.
5. Оценка полученного текста на предмет его осмысленности и соответствия языковым признакам.

Такой метод, несмотря на свою наглядность, имеет ряд очевидных ограничений. Наиболее существенным из них является фактическая невозможность ручного перебора всех комбинаций

(без применения автоматизации) при длине ключа, превышающей 8–9 символов, особенно без автоматизации и при ручном анализе.

Частично эту проблему можно решить за счет эвристик, например, перебора наиболее вероятных длин ключей или предположений о начале открытого текста. Однако такие эвристики нередко сводятся к догадкам и не всегда способствуют развитию аналитического мышления у студентов. Более того, в ряде методических пособий наблюдается тенденция предлагать учащимся лишь укороченные варианты задач, в которых заранее известна длина ключа или задан упрощенный алфавит, что снижает обучающую ценность таких заданий.

Подробные методики криптоанализа транспозиционных шифров, включая вертикальные, изложены в классических работах У. Ф. Фридмена [7] и Х. Ф. Гейнс [8], которые содержат как теоретические основы, так и практические приемы анализа.

Таким образом, традиционный метод взлома шифра вертикальной перестановки, основанный на последовательном переборе всех возможных вариантов, демонстрирует низкую эффективность при увеличении длины ключа и не позволяет обучающимся в полной мере реализовать принципы анализа вероятностных структур и языковых закономерностей, применимых в реальных условиях криптоанализа.

Для преодоления указанных ограничений представляется целесообразным использовать более интеллектуализированные методы, основанные, например, на статистическом анализе текста. Одним из таких методов является частотный криптоанализ, ориентированный на проверку вероятности появления конкретных триграмм или других статистических единиц в полученном шифртексте.

МОДИФИЦИРОВАННЫЙ МЕТОД ЧАСТОТНОГО АНАЛИЗА

В рамках настоящего исследования предлагается модифицированный метод криптоанализа классического шифра вертикальной перестановки, основанный на использовании частотной информации о триграммах естественного языка. В отличие от традиционного перебора всех возможных вариантов перестановки столбцов, предлагаемый подход позволяет существенно сократить количество рассматриваемых комбинаций путем предварительного исключения тех, которые заведомо не могут соответствовать реальному языковому тексту.

Основная эвристическая посылка метода заключается в следующем: если при сопоставлении трех столбцов шифртекста в любом порядке образуются буквосочетания, не встречающиеся в языке (например, "ЙЩЪ", "ФФЧ" и подобные), то данное расположение этих столбцов можно считать лингвистически невозможным и исключить его из дальнейшего анализа. Таким образом, криптоаналитику нет необходимости проверять все $n!$ перестановок (где n – длина ключа), поскольку большая их часть будет содержать фонетически и статистически маловероятные триграммы.

Таким образом, модифицированный алгоритм частотного анализа шифра вертикальной перестановки состоит из следующих этапов:

1. Формирование триграммной базы.

На основе анализа литературных текстов на русском языке (в данном случае использовался рассказ И. С. Тургенева «Муму») составляется список возможных трехбуквенных сочетаний (триграмм), реально встречающихся в языке.

– алфавит включает 34 символа, включая пробел: {_, А, Б, …, Я};

– весь анализируемый текст предварительно нормализуется: удаляются знаки препинания, все символы приводятся к верхнему регистру, множественные пробелы заменяются на одиночные;

– далее создается трехмерный массив размером 34×34×34 (рис. 1), где каждая ячейка отражает факт наличия соответствующей триграммы в тексте. Пример процесса представлен на рисунках 2–3.

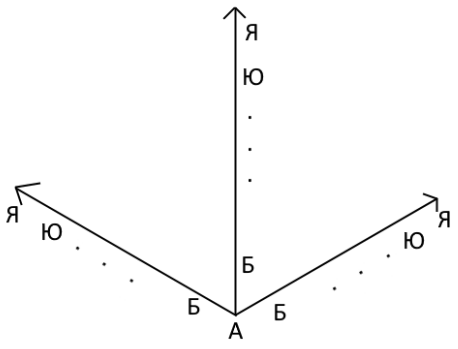


Рисунок 1 – Трехмерный массив

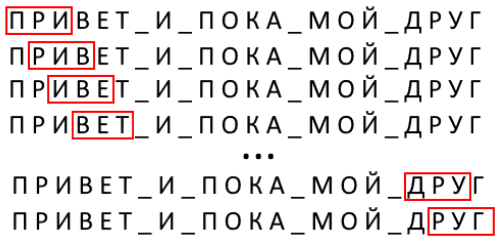


Рисунок 2 – Процесс перебора триграмм в тексте «ПРИВЕТ_И_ПОКА_МОЙ_ДРУГ»

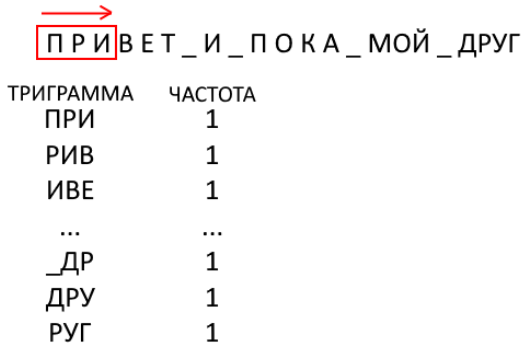


Рисунок 3 – Результат подсчета триграмм

В результате анализа вышеупомянутого литературного произведения были выделены 6 126 встречающихся триграмм, а 33 178 сочетаний не встретилось ни разу. Поскольку 33 178 ≫ 6 126, целесообразно ограничить анализ только допустимыми триграммами.

2. Определение длины ключа.

Предполагается длина ключа, например 10, и на основе этой длины шифртекст преобразуется в таблицу, где каждая строка содержит по 10 символов.

3. Генерация допустимых сочетаний столбцов.

Вместо полного перебора всех 10! = 3 628 800 вариантов, анализируются только те сочетания трех столбцов (всего P(10,3) = 720), которые при составлении текстов дают допустимые триграммы;

Если в определенной комбинации столбцов обнаруживаются запрещенные триграммы, то вся перестановка, включающая эти столбцы в таком порядке, исключается из дальнейшего анализа.

4. Формирование кандидатов на дешифровку.

Из всех допустимых сочетаний трех столбцов формируются полные перестановки, включающие только совместимые фрагменты. Далее они применяются к шифртексту, и полученные варианты оцениваются по частоте появления валидных триграмм.

5. Ранжирование вариантов.

Каждой перестановке присваивается вес, соответствующий количеству найденных допустимых триграмм, что позволяет выбрать наиболее вероятные расшифровки.

Применение частотного фильтра на ранних этапах существенно снижает вычислительные затраты. При длине ключа в 10 символов рассматривается не более 720 триграммных сочетаний, вместо миллионов полных перестановок. Это делает предложенный метод особенно удобным для образовательных целей, а также для ручного анализа или анализа с применением ограниченных вычислительных ресурсов.

ПРИМЕР РЕАЛИЗАЦИИ МЕТОДА

Для демонстрации эффективности предложенного метода частотного анализа вертикального шифра перестановки рассмотрим конкретный пример. Предположим, что известен зашифрованный текст: **«ПГЯДОЛИЬКАШ_КНСМ_ЕЕН_РТОАРСТЗААТЯ_ЕСЗОСЛЛ_»**, и предполагается, что использован классический шифр вертикальной перестановки.

1. Определение длины ключа.

Первоначально необходимо установить предполагаемую длину ключа. Поскольку длина шифртекста составляет 42 символа, потенциальными длинами ключа являются делители этого числа: 2, 3, 6, 7, 14, 21, 42. Для дальнейшего анализа принимаем длину ключа равной 6, так как она позволяет получить сбалансированную таблицу (матрицу) из 6 столбцов.

2. Построение шифротаблицы.

Текст записывается по строкам в таблицу из 6 столбцов, заполняемую слева направо и сверху вниз (табл. 1).

Таблица 1 – Заполнение таблицы (I этап)

1	2	3	4	5	6
П	Г	Я	Д	О	Л
И	Ь	К	А	Ш	–
К	Н	С	М	–	Е
Е	Н	-	Р	Т	О
А	Р	С	Т	З	А
А	Т	Я	–	Е	С
З	О	С	Л	Л	–

3. Анализ возможных сочетаний столбцов.

Следующий шаг – анализ возможных комбинаций трех столбцов, которые потенциально могут располагаться рядом в исходном (открытом) тексте. Общее число таких размещений составляет:

$$A_6^3 = \frac{6!}{(6-3)!} = 120.$$

Для каждой тройки столбцов проверяется вероятность появления соответствующих триграмм на основе ранее составленной статистики по русскому языку. Например, рассмотрим гипотезу, что рядом находятся столбцы 1, 2 и 4. Запишем их подряд и проанализируем частотность появляющихся триграмм (табл. 2).

Таблица 2 – Смена столбцов 3 и 4

1	2	4	3	5	6
П	Г	Д	Я	О	Л
И	Ь	А	К	Ш	–
К	Н	М	С	–	Е
Е	Н	Р	–	Т	О
А	Р	Т	С	З	А
А	Т	-	Я	Е	С
З	О	С	С	Л	–

Первая полученная триграмма – «ПГД» – отсутствует в корпусе возможных триграмм, вероятность ее появления равна нулю. Следовательно, данное сочетание отбрасывается.

Затем рассматривается другая последовательность, например, столбцы 1, 4 и 2, формирующие триграмму «ПОГ» (табл. 3). Эта комбинация имеет ненулевую частоту появления (в исследованном корпусе – 57 раз), что делает ее допустимой.

Таблица 3 – Смена столбцов

1	5	2	3	4	6
П	О	Г	Я	Д	Л
И	Ш	Ь	К	А	–
К	–	Н	С	М	Е
Е	Т	Н	–	Р	О
А	З	Р	С	Т	А
А	Е	Т	Я	–	С
З	Л	О	С	Л	–

Аналогичным образом проверяются все остальные триграммы, сформированные из этой последовательности столбцов. Если все триграммы оказываются валидными, можно сделать вывод о высокой вероятности корректности данной комбинации (табл. 4).

Таблица 4 – Смена столбцов и проверка валидности триграмм

1	5	2	3	4	6
П	О	Г	Я	Д	Л
И	Ш	Ь	К	А	–
К	-	Н	С	М	Е
Е	Т	Н	–	Р	О
А	З	Р	С	Т	А
А	Е	Т	Я	–	С
З	Л	О	С	Л	–

Таким способом подтверждается первая группа смежных столбцов. Однако их точное положение в шифротаблице (в начале, середине или конце) на данном этапе не установлено.

4. Построение полной перестановки.

Далее продолжается перебор оставшихся столбцов, например анализируется последовательность 5, 2 и 3. Если одна из триграмм, например «ОГЯ», имеет нулевую частотность, данное сочетание также исключается. Пробуются альтернативные комбинации, такие как 5, 2 и 6 и т. д., до тех пор пока не будут определены все допустимые комбинации столбцов на основании анализа частотности триграмм (рис. 4).

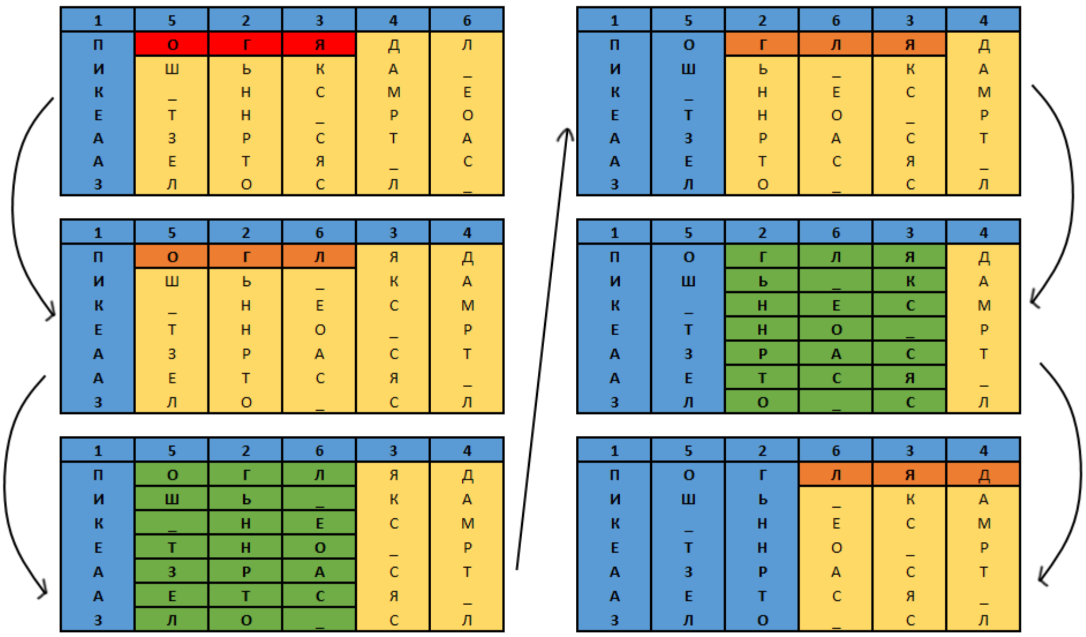


Рисунок 4 – Перебор столбцов

5. Восстановление открытого текста.

После установления порядка столбцов формируется полная таблица с переставленными столбцами, из которой по строкам восстанавливается исходный текст (табл. 5).

Таблица 5 – Окончательный вид таблицы с правильным порядком столбцов

1	5	2	6	3	4
П	О	Г	Л	Я	Д
И	Ш	Ь	–	К	А
К	–	Н	Е	С	М
Е	Т	Н	О	–	Р
А	З	Р	А	С	Т
А	Е	Т	С	Я	–
З	Л	О	–	С	Л

В рамках данного примера итоговый расшифрованный текст выглядит следующим образом: «ПОГЛЯДИШЬ_КАК_НЕСМЕТНО_РАЗРАСТАЕТСЯ_ЗЛО_СЛ...».

Таким образом, предложенный модифицированный метод частотного криптоанализа шифров перестановки позволяет значительно сократить объем вычислений по сравнению с полным перебором всех возможных вариантов перестановок.

ЗАКЛЮЧЕНИЕ

В ходе исследования был разработан и апробирован частотный метод криптоанализа шифра вертикальной перестановки, основанный на анализе допустимых триграмм в естественном (русском) языке. Метод позволяет существенно ограничить пространство перебора возможных ключей, исключая комбинации, приводящие к маловероятным или невозможным с точки зрения языка триграммам. Проведенный эксперимент показал, что предложенный подход способен эффективно восстанавливать структуру открытого текста даже в условиях отсутствия ключа. Использование частотной информации повышает точность анализа и снижает вычислительные затраты, что особенно актуально при работе с большими объемами зашифрованных данных.

Особую значимость данный подход приобретает в образовательной практике. Дисциплина «Криптографические методы защиты информации» играет ключевую роль в формировании профессиональных компетенций будущих специалистов по информационной безопасности. В процессе изучения данной дисциплины студенты знакомятся не только с современными алгоритмами, но и с историческими шифрами, анализ которых позволяет глубже понять фундаментальные принципы криптографии. Критический разбор уязвимостей, включая методы криптоанализа, развивает у обучающихся навыки логического мышления и системного криптоанализа.

Таким образом, частотный криптоанализ может быть рекомендован не только как вспомогательный или основной метод анализа шифров перестановки, но и как важный инструмент для формирования аналитического мышления в рамках подготовки специалистов по защите информации. Его применение в учебном процессе позволяет эффективно сочетать теоретические знания с практическими навыками, формируя у студентов устойчивое понимание принципов информационной безопасности.

Список источников

1. Шнайер, Б. Прикладная криптография: протоколы, алгоритмы и исходные тексты на языке С / Б. Шнайер. – Москва : Диалектика-Вильямс, 2019.

2. Тилборг ван, Х. К. А. Основы криптологии. Профессиональное руководство и интерактивный учебник / Х. К. А. ван Тилборг. – Москва : Мир, 2006. – 471 с. – ISBN 5-03-003639-3.

3. Введение в криптографию / под общей ред. В. В. Ященко. – Санкт-Петербург : Питер, 2001. – 288 с. – ISBN 5-318-00443-1

4. Stallings, W. Cryptography and Network Security: Principles and Practice / W. Stallings. – Pearson, 2022.

5. Басалова, Г. В. Основы криптографии : учебное пособие / Г. В. Басалова. – 2-е изд. – Москва : ИНТУИТ, 2016. – 282 с. // Лань : электронно-библиотечная система. – URL: https://e.lanbook.com/book/100302 (дата обращения: 06.07.2025).

6. Рябко, Б. Я. Основы современной криптографии и стеганографии : монография / Б. Я. Рябко, А. Н. Фионов. – Москва : Горячая линия-Телеком, 2011. – 232 с. – ISBN 978-5-9912-0150-6 // Лань : электронно-библиотечная система. – URL: https://e.lanbook.com/book/5192 (дата обращения: 07.07.2025).

7. Friedman, W. F. Military Cryptanalysis. Part IV: Transposition and Fractionating Systems / W. F. Friedman.

8. Gaines, H. F. Cryptanalysis: A Study of Ciphers and Their Solution / H. F. Gaines. – 1989. – 256 p.

References

1. Schneier, B. *Applied cryptography: protocols, algorithms and source texts in the language of S.* Moscow, Dialectics-Williams Publ., 2019 (In Russ.).

2. Tilborg van, H. K. A. *Fundamentals of Cryptology. Professional guide and interactive textbook*. Moscow, Mir Publ., 2006. 471 p. ISBN 5-03-003639-3 (In Russ.).

3. *Introduction to cryptography.* Under the general editorship of V. V. Yashchenko. St. Petersburg, Peter Publ., 2001. 288 p. ISBN 5-318-00443-1 (In Russ.).

4. Stallings, W. *Cryptography and Network Security: Principles and Practice*. Pearson, 2022.

5. Basalova, G. V. Fundamentals of cryptography: a textbook. 2nd ed. Moscow, INTUIT, 2016. 282 p. *Lan : electronic library system*. Available at: https://e.lanbook.com/book/100302 (accessed 06.07.2025) (In Russ.).

6. Ryabko, B. Ya., Fionov, A. N. Fundamentals of modern cryptography and steganography : a monograph. Moscow, Goryachaya Liniya-Telecom Publ., 2011. 232 p. ill. ISBN 978-5-9912-0150-6. *Lan : electronic library system.* Available at: https://e.lanbook.com/book/5192 (accessed 07.07.2025) (In Russ.).

7. Friedman, W. F. *Military Cryptanalysis. Part IV: Transposition and Fractionating Systems*.

8. Gaines, H. F. *Cryptanalysis: A Study of Ciphers and Their Solution*, 1989. 256 p.

Статья поступила в редакцию 13.08.2025; одобрена после рецензирования 22.09.2025; принята к публикации 26.09.2025.

The article was submitted 13.08.2025; approved after reviewing 22.09.2025; accepted for publication 26.09.2025.

УДК 004.032.26

ИДЕНТИФИКАЦИЯ БОТОВ В СОЦИАЛЬНЫХ СЕТЯХ МЕТОДАМИ АНАЛИЗА ТЕКСТА И ПОВЕДЕНИЯ

Частикова Вера Аркадьевна, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
кандидат технических наук, доцент, ORCID: 0000-0003-2372-8275, e-mail: chastikova_va@mail.ru
Козачёк Константин Валерьевич, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
аспирант, ORCID: 0009-0002-9945-2113, e-mail: Koza4ek.Konstantin@yandex.ru
Крамарь Александр Викторович, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
студент, ORCID: 0009-0002-7957-2767, e-mail: alex.ign.231103@gmail.com
Маматов Ян Муратович, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
студент, ORCID: 0009-0000-4946-965X, e-mail: mamatov.yan2003@gmail.com

В статье рассматривается задача автоматического выявления ботов в социальных сетях. Предложен подход, основанный на совместном анализе текстовой информации и особенностей поведения пользователей. Разработанная нейросетевая модель анализирует лингвистические характеристики сообщений и выявляет аномалии в активности аккаунтов. Экспериментальные результаты демонстрируют эффективность подхода для решения задачи классификации. Полученное решение может быть использовано для автоматизации процессов модерации в социальных платформах. Применение подобных систем способствует повышению качества взаимодействия пользователей в цифровом пространстве.

Ключевые слова: нейросетевая модель, идентификация ботов, социальная сеть, анализ текста, поведенческий паттерн, ruBERT, машинное обучение, классификация пользователей

IDENTIFICATION OF BOTS IN SOCIAL NETWORKS BY TEXT AND BEHAVIOR ANALYSIS METHODS

Chastikova Vera A., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,
Cand. Sci. (Engineering), Associate Professor, ORCID: 0000-0003-2372-8275, e-mail: chastikova_va @mail.ru
Kozachek Konstantin V., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,
graduate student, ORCID: 0009-0002-9945-2113, e-mail: Koza4ek.Konstantin@yandex.ru
Kramar Alexander V., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,
student, ORCID: 0009-0002-7957-2767, e-mail: alex.ign.231103@gmail.com
Mamatov Yan M., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,
student, ORCID: 0009-0000-4946-965X, e-mail: mamatov.yan2003@gmail.com

This article examines the problem of automatic detection of bots in social networks. A method based on a joint analysis of textual information and user behavior features is proposed. The developed neural network model analyzes the linguistic characteristics of messages and identifies anomalies in account activity. The experimental results demonstrate the effectiveness of the approach for solving the classification problem. The resulting solution can be used to automate moderation processes on social platforms. The use of such systems helps to improve the quality of user interaction in a digital environment.

Keywords: neural network model, bot detection, social network, text analysis, behavioral pattern, ruBERT, machine learning, users classification

Графическая аннотация (Graphical annotation)

ВВЕДЕНИЕ

Социальные сети давно перестали быть просто площадкой для общения – они влияют на новости, политику, бизнес и настроение общества. На фоне этого все больше внимания привлекают боты – автоматические аккаунты, которые ведут себя как люди, но используются для манипуляции мнением, распространения ложной информации или создания искусственной активности. Визуально они могут казаться обычными пользователями, но за ними нет реального человека.

Сейчас большинство систем, которые пытаются находить ботов, обращают внимание в основном на технические детали – вроде количества лайков, скорости публикаций или IP-адресов. Но такие признаки легко подделать. Если же смотреть не только на цифры, а еще и на то, как именно пишется текст, какие слова используются, как строятся фразы, а также когда и как часто человек проявляет активность, можно заметить отличия, которые сложнее подделать. Такой подход позволяет точнее понимать, с кем мы имеем дело – с реальным пользователем или машиной, которая старается выглядеть как человек.

Разработанная система может быть полезной в самых разных ситуациях – от чистки фейковых аккаунтов до выявления вбросов в инфополе. Ее можно применять как в рамках соцсетей, так и на платформах с пользовательским контентом. Кроме того, работа помогает глубже понять, чем отличается поведение настоящих людей от автоматизированных систем и какие мелочи выдают ботов, даже если с виду они выглядят как обычные пользователи.

С начала 2020-х гг. Россия столкнулась с беспрецедентным ростом информационной активности в цифровом пространстве, где ключевую роль играют автоматизированные аккаунты – боты. Их активность связана как с внутренними, так и внешними событиями: пандемией, электоральными циклами, санкциями и геополитическими кризисами. Важно отметить, что Россия является не только источником, но и объектом дезинформационных кампаний, в том числе координируемых из-за рубежа. Эти процессы вынуждают российские компании, ведомства и научные организации развивать более продвинутые технологии детекции и нейтрализации автоматических агентов в сети.

ПРИМЕРЫ УЩЕРБА ОТ БОТ-КАМПАНИЙ

Социальные боты приносят реальный ущерб и гражданам, и государству. Классический пример – финансовые мошенничества через чат-боты. Например, в марте 2025 г. МВД РФ зафиксировало новую схему: в мессенджере Telegram появился фейковый бот «Почты России», предлагающий «найти посылку». Пользователь вводил данные, а мошенники похищали его деньги [1]. Аналогичные телеграм- и соцсетевые боты-мошенники обещают подарки или услуги (например, «бесплатная подписка Telegram Premium») и выманивают пароли и деньги у наивных юзеров [1]. В масштабах страны ущерб от IT-мошенников (в том числе через ботов) огромен: за 11 мес. 2024 г. россияне

потеряли свыше 168 млрд руб. (страховые компании и МВД фиксируют более 700 тыс. «информационных» преступлений) [2]. Соцсети и властные структуры пытаются пресечь фейковые вбросы: при расследовании коронавирусного фейка Group-IB сотрудничал с администрацией «ВКонтакте» и Facebook и удалил атаку «аудиопранка» практически полностью [3]. Тем не менее сотни тысяч «фейковых» сообщений (о COVID, мобилизации, политике и т. д.) блокируются вручную Роскомнадзором и платформами только после широкого разбирательства. Числа украденных средств приводятся разными ведомствами отчасти разрозненно, но ясно, что боты-фейки напрямую приводят к значительным финансовым потерям и подрыву доверия к официальной информации [2].

УЧАСТИЕ РОССИЙСКИХ КОМПАНИЙ В БОРЬБЕ С БОТАМИ

В России борьбу с ботами ведут и коммерческие IT-компании, и научные организации. «Лаборатория Касперского» регулярно публикует разборы мошеннических схем с ботами: её специалисты, например, подробно описали фейковые телеграм-боты «Почты России» и показали, насколько они «продвинутые» и неотличимы от настоящих [4]. Аналитики Kaspersky предупреждают пользователей и поддерживают правоохранителей, создавая базы известных фейковых ботов. Group-IB (частная антивирусная компания) вместе с МВД РФ мониторит бот-атаки: у нее есть сервис сбора данных о ботах, передаваемый правоохранителям и администраторам сетей [3]. Этот канал сотрудничества помогает «ВКонтакте», Facebook и Telegram быстрее блокировать дезинформацию и мошенников. Официально «ВКонтакте» также заявляет о борьбе с ботами: помимо ручного модераторского тега система использует алгоритмы поведения (частота постинга, новые аккаунты без фото и др.) и машинное обучение для массового удаления подозрительных аккаунтов.

ИССЛЕДОВАНИЯ В НАУЧНОМ СООБЩЕСТВЕ

Исследователи ВШЭ (напр. Д. Самохвалов) собрали крупные выборки «злонамеренных» и «чистых» аккаунтов VK и применили CatBoost на текстовых и графовых признаках, добившись AUC ≈ 0,91 [5]. Ученые СПбФИЦ РАН (А. Чечулин и М. Коломеетов) предложили подход «купить» доступ к ботам и «Тест Тьюринга» среди пользователей, чтобы обучать модели на признаках из графа взаимодействий и текстовых сообщений [7]. Партнеры банков и телекомов (например, Сбер, МТС) используют технологии NLP и поведенческого анализа для защиты клиентов от фишинга и ботов-мошенников, хотя публикуют по этой теме меньше открытой информации. В академии также задействованы ИТ-факультеты МГУ, МФТИ, РАНХиГС, где исследуют социальные сети с помощью нейронных сетей, статистики и сетевого анализа. В совокупности российские компании и университеты используют сочетание классических методов (правила, частотные фильтры, простой ML) и современных – глубокое обучение (нейросети на текстах), графовые алгоритмы и инструментальные ML-платформы.

Исследования в научном сообществе охватывают широкий спектр методов распознавания ботов. Во-первых, это анализ текста и естественного языка. Например, Ng и Carley (2023, Carnegie Mellon) выявили, что боты часто используют шаблонные лингвистические приемы: много хэштегов, положительные слова, в то время как люди чаще вступают в диалог, отвечают на чужие комментарии [8]. Подобные «лингвистические маркеры» активно используются в ботодетекторе Botometer (Indiana University) [10], где тексты аккаунтов векторизуются и подаются на обученные классификаторы. Во-вторых, исследователи изучают поведение и активность: расчет средней частоты постинга, временных паттернов (боты могут писать круглосуточно или слишком регулярно), долю ретвитов/репостов и т. п. Эта информация часто проявляет «звездную» сетевую структуру ботов, тогда как обычные пользователи образуют иерархии общения [8]. Наконец, анализы структуры графа друзей и взаимодействий показали, что в графовых признаках (встроенных эмбеддингах узлов) содержится много информации о ботах. Модель, основанная на узловых и структурных эмбеддингах социальных графов, превзошла по эффективности подходы, опирающиеся только на текст или метаданные [9]. В целом современные ботодетекторы комбинируют текстовые характеристики (семантика, стилистика сообщений), поведенческие модели (как часто и как быстро публикуется контент) и сетевые признаки (структура связей). Большинство проверенных методик показывают высокую точность (AUC 0,8–0,95) в закрытых тестах, хотя в «полевых» условиях эффективность может снижаться из-за постоянной эволюции ботов.

СРАВНЕНИЕ С ПОДХОДОМ НА RUBERT

Подход, реализованный в данной работе, фокусируется на анализе текстового контента с помощью модели RuBERT (российский BERT). Это дает глубокое понимание смысла и стилистики сообщений на русском языке. В отличие от простых методов (мешок слов (Bag-of-Words) или TF-IDF), RuBERT улавливает контекстуальные связи между словами и может различать, например, слишком «безликий» или шаблонный слог бота и естественную речь человека. Такой подход хорошо улавливает «лингвистические сигнатуры» ботов (напомним, что боты часто применяют одни и те же паттерны слов и эмоций [8]). Он сопоставим с иностранными работами, где применяются трансформеры для классификации фейковых новостей и деанонимизации (например, BERT-классификаторы для распознавания политических троллей).

Однако подход на основе только текста имеет ограничения. В отличие от многомодальных схем (как у Самохвалова или Dehghan), он не учитывает граф друзей и временные паттерны. Если бот умело маскируется контентом и пишет «человеческие» тексты (например, его сообщения тщательно генерирует нейросеть), то RuBERT может не заметить его «злонамеренности» без дополнительных признаков. Напротив, графовые методы улавливают типичную изоляцию или повторяемость связей у ботов [9], а методы по частотному анализу фиксируют аномально высокую активность. В идеале для максимальной точности текстовые модели (RuBERT) стоит комбинировать с признаками поведения: например, выдавать более высокую вероятность бота, если аккаунт помимо подозрительной лексики дополнительно публикует сотни постов в день или дружит с сотнями других ботов.

В сумме метод на RuBERT обеспечивает мощную основу для выявления ботов по содержанию сообщений (что выделяет его среди многих существующих решений). В опубликованных исследованиях сообщалось о высокой эффективности именно таких NLP-подходов при классификации угроз (заявляют AUC ~ 0,9 [5]). Лучшие результаты обычно получают гибридные системы. Таким образом, подход с RuBERT следует рассматривать как сильное дополнение к существующим методам: он особенно эффективен против ботов с шаблонным текстом, но для максимально надежной работы нуждается в учете и других источниках информации (граф друзей, временной анализ, метаданные).

ОПИСАНИЕ РЕАЛИЗОВАННОГО ПОДХОДА

Языковые модели, основанные на архитектуре трансформеров (вроде BERT или его русскоязычной версии RuBERT), хорошо справляются с анализом текста благодаря способности учитывать контекст слов. В отличие от традиционных методов, где слова обрабатываются по отдельности, трансформеры видят предложение целиком и могут различать даже тонкие смысловые оттенки. Это особенно важно, когда речь идет о различии между настоящим пользователем и ботом – ведь боты, как правило, пишут шаблонно, используют повторяющиеся паттерны и избегают сложных конструкций.

Прежде чем перейти к использованию языковых моделей, в задаче анализа текста долгое время использовались более простые статистические подходы. Один из них – метод TF-IDF (term frequency – inverse document frequency). Суть TF-IDF в том, чтобы определить значимость слова в контексте всего корпуса текстов. Слова, которые встречаются часто в одном документе, но редко – в других, получают высокий вес. Это позволяет отсеять общие слова вроде «и», «в», «на» и выделить информативные термины. Однако TF-IDF работает на уровне слов и не учитывает порядок слов или контекст.

Еще один распространенный метод – «мешок слов» (Bag of Words, BoW). Он совсем не смотрит на порядок слов: предложение «бот активен ночью» и «ночью активен бот» будут для модели абсолютно одинаковыми. Слова просто превращаются в векторы, где каждому слову соответствует число вхождений. Простой, но довольно грубый способ, который часто использовался до появления нейросетей.

На смену этим подходам пришли методы глубокого обучения [6]. Особенно большую популярность получили предобученные языковые модели – они уже прошли этап обучения на огромных массивах текста, что позволяет им «понимать» структуру языка и адаптироваться под конкретную задачу, например – детекцию ботов. В данном исследовании была использована модель cointegrated/rubert-tiny2 – компактная, но довольно мощная версия RuBERT, оптимизированная для работы с русским языком. Ее преимущество в том, что она позволяет быстро и точно классифицировать текст, при этом не требует большого количества вычислительных ресурсов.

Перед тем как подать текст на вход модели, он проходит этап предобработки. Это включает приведение к нижнему регистру, удаление ссылок, HTML-тегов, лишних символов и пробелов. Далее текст превращается в токены – числовые представления слов, которые подаются в модель. Обучение происходит с учетом весов классов, чтобы избежать перекоса в пользу одного из классов, если данных «ботов» или «не ботов» больше.

В результате модель успешно обучается различать, является ли текст, опубликованный в социальной сети, ботогенерированным. На тестовой выборке точность достигла 95–97 %, что указывает на высокий потенциал таких решений в реальной практике.

Таким образом, можно сделать вывод, что языковые модели – это не просто инструмент перевода или поиска, а полноценный механизм анализа поведения в цифровом пространстве. Совместно с методами машинного обучения они дают возможность выявления ботов с высокой точностью, помогая бороться с дезинформацией и обеспечивать безопасность информационного пространства.

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА

Модель обучалась на наборе данных с использованием 8 эпох. Пример датасета, используемого для обучения нейросети, приведен на рисунке 1.

```
{
    "id": 241,
    "text": "Ваш код подтверждения: 123456",
    "source": "Telegram",
    "label": 1
},
{
    "id": 81,
    "text": "Кто-нибудь может порекомендовать хорошую книгу?",
    "source": "Odnoklassniki",
    "label": 0
},
{
    "id": 75,
    "text": "Кто-нибудь идёт на концерт в субботу?",
    "source": "Odnoklassniki",
    "label": 0
},
{
    "id": 206,
    "text": "Нажмите здесь, чтобы получить дополнительную информацию.",
    "source": "VK",
    "label": 1
},
{
    "id": 199,
    "text": "Пожалуйста, подождите, мы обрабатываем ваш запрос.",
    "source": "Telegram",
    "label": 1
},
```


Рисунок 1 – Пример данных для обучения модели

За время обучения наблюдался рост точности на обучающих данных, а также улучшение значений метрик на валидационных данных. Графики точности и потерь на обучающем и валидационном наборах данных приведены на рисунке 2.

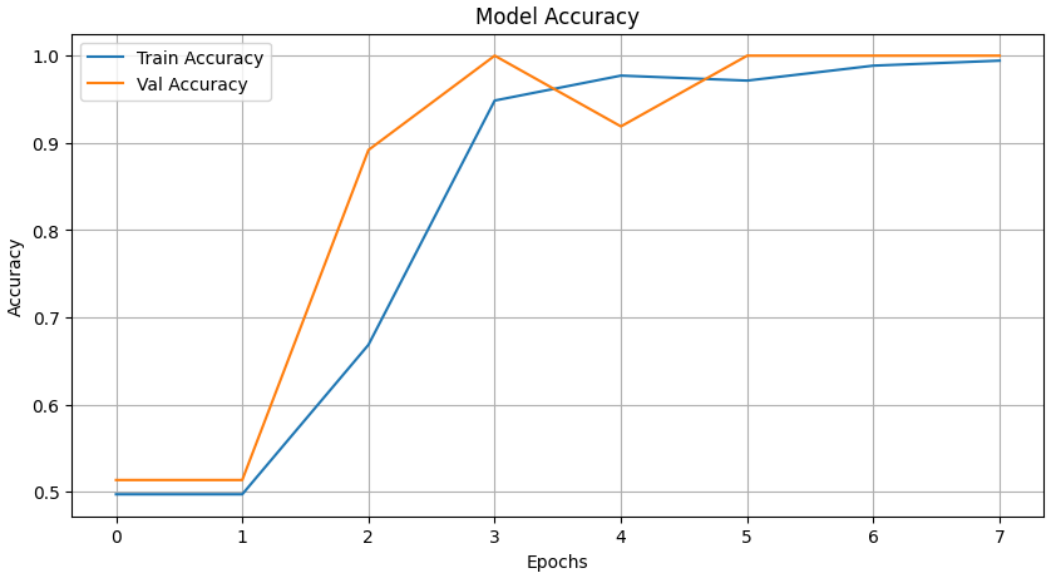


Рисунок 2 – Точность модели на обучающих и валидационных данных

На этом графике (рис. 2) показано, как точность модели изменялась по мере обучения. Можно заметить, что точность на обучающих данных стабильно возрастала с каждой эпохой, в то время как на валидационных данных точность также увеличивалась и достигала стабильного уровня. Из графика видно, что модель постепенно улучшала свои результаты на обоих наборах данных. Особенно это заметно на 4-й эпохе, когда точность на валидации стабилизировалась, что свидетельствует о хорошем обучении модели и предотвращении переобучения.

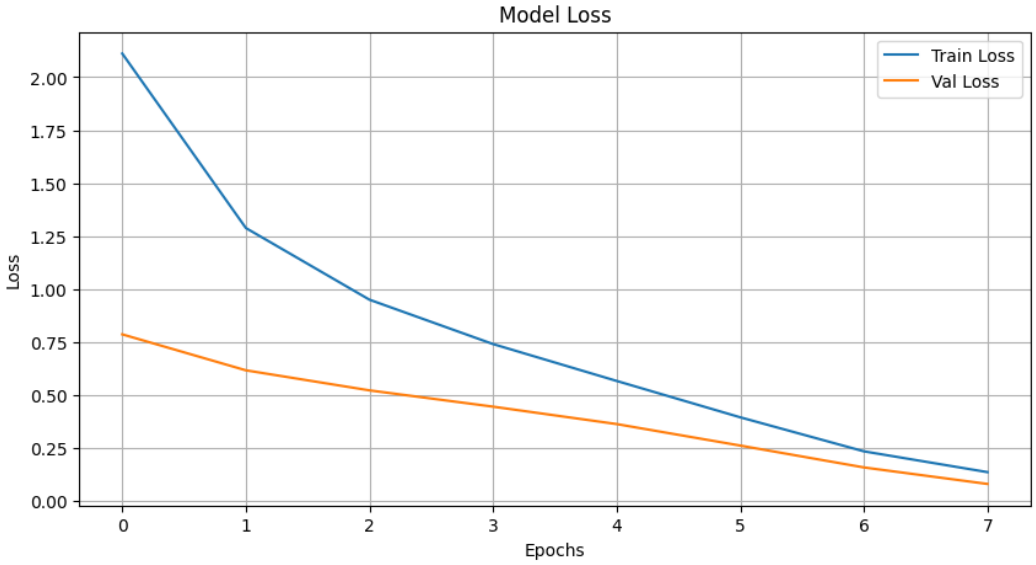


Рисунок 3 – Потери модели на обучающих и валидационных данных

На графике (рис. 3) представлены потери модели на обучающих и валидационных данных на протяжении всего обучения. Можно увидеть, как потери на обучающих данных постепенно уменьшались, что указывает на повышение эффективности модели, в то время как на валидационных данных потери стабилизировались и стали минимальными. Как видно из графика, потери на обучающих данных постепенно снижались, а на валидации оставались стабильными после 3-й эпохи. Это указывает на то, что модель хорошо обучалась и был достигнут оптимальный уровень производительности.

После завершения обучения модель была протестирована на отдельной тестовой выборке. Результаты предсказаний модели показали высокие значения точности и AUC. Модель демонстрирует способность точно классифицировать тексты на категории «бот» и «не бот».

ПРИМЕР РАБОТЫ ПРОГРАММЫ

Для демонстрации работы обученной модели на реальных данных были приведены несколько примеров классификации текстов, введенных пользователем через консоль. Модель получает на вход текст и выдает вероятность того, что это сообщение отправлено ботом. В зависимости от вероятности, программа классифицирует текст как «бот» или «не бот».

Рассмотрим два примера:

1. Пользователь вводит сообщение: «Я люблю гулять с собакой по утрам».

Модель оценивает вероятность того, что это сообщение отправлено ботом как 0,4638, что указывает на низкую вероятность. Таким образом, модель классифицирует это сообщение как «не бот».

2. Пользователь вводит сообщение: «Твой код подтверждения: 345678».

В данном случае вероятность того, что сообщение отправлено ботом, составляет 0,7450, что достаточно высоко. Модель классифицирует это сообщение как «бот».

Подобные примеры наглядно демонстрируют, как модель может различать текстовые сообщения и присваивать им соответствующие категории. Для лучшего понимания ниже приведен скриншот с результатом работы программы (рис. 4).

```
Your message: Я люблю гулять с собакой по утрам.
1/1 [==============================] - 0s 47ms/step
Bot probability: 0.4638

Your message: Ваш код подтверждения: 123456.
1/1 [==============================] - 0s 45ms/step
Bot probability: 0.7450
```

Рисунок 4 – Пример вывода программы

ЗАКЛЮЧЕНИЕ

В ходе проведенных исследований была разработана и обучена модель на основе предобученного трансформера для задачи бинарной классификации сообщений на категории «бот» и «не бот». Модель показала высокую точность на валидационной выборке и адекватные результаты на тестовых данных, что подтверждается как числовыми метриками, так и визуализацией процесса обучения.

Также была реализована CLI-версия, позволяющая в реальном времени проверять тексты и получать вероятностную оценку принадлежности к классу. Такой подход делает модель не только исследовательским инструментом, но и потенциально применимой в практических задачах, связанных с фильтрацией контента или анализом пользовательской активности. Полученные результаты показывают, что даже компактные модели могут успешно справляться с задачей определения ботов при корректной предобработке и настройке гиперпараметров.

Список источников

1. Розанова, А. В. Мошенники начали похищать деньги через фейковый бот «Почты России» / А. В. Розанова // РБК Life. – 2025. – URL: https://www.rbc.ru/life/news/67c7ef979a79472746515cc2 (дата обращения: 20.04.2025).

2. Бородовский, Д. А. За 2024 год мошенники украли у россиян 168 млрд рублей / Д. А. Бородовский // Hi-Tech Mail. – 2025. – URL: https://hi-tech.mail.ru/news/120371-za-2024-god-moshenniki-ukrali-u-rossiyan-168-mlrd-rublej/ (дата обращения: 20.04.2025).

3. Анисимова, Н. А. Данные распространявших фейк о коронавирусе аккаунтов передали силовикам / Н. А. Анисимова // РБК.РУ. – 2020. – URL: https://www.rbc.ru/rbcfreenews/5e5fbff99a7947c60350de28 (дата обращения: 20.04.2025).

4. Каминский, С. Н. Мошенничество с ботом «Почты России» в Telegram / С. Н. Каминский // kaspersky daily. – 2025. – URL: https://www.kaspersky.ru/blog/russian-post-telegram-scam/39383/ (дата обращения: 20.04.2025).

5. Самохвалов, Д. А. Machine Learning-Based malicious users detection in the VKontakte social network / Д. А. Самохвалов // Semantic Scholar. 2020. URL: https://pdfs.semanticscholar.org/5db8/3d3e88b653e1d2242d547420d99f06a09228.pdf (дата обращения: 20.04.2025).

6. Частикова, В. А. Технологии искусственного интеллекта в информационной безопасности : монография / В. А. Частикова, С. А. Жерлицын, С. А. Митюгов. – Краснодар : Изд-во КубГТУ, 2024. – 315 с.

7. Чечулин, А. А. Подход к обнаружению вредоносных ботов в социальной сети Вконтакте и оценка их параметров / А. А. Чечулин, М. В. Коломеец // Труды учебных заведений связи. – 2024. – Т. 10, № 2. – С. 92–101. – DOI: 10.31854/1813-324X-2024-10-2-92-101. – URL: https://www.researchgate.net/publication/380418227 (дата обращения: 20.04.2025).

8. Ng, L. What is a Social Media Bot? A Global Comparison of Bot and Human Characteristics / L. Ng, K. Carley // arxiv.org. – 2025. – URL: https://arxiv.org/html/2501.00855v1 (дата обращения: 20.04.2025).

9. Denghan, A. Suita K. Detecting bots in social-networks using node and structural embeddings / A. Denghan, K. Suita // Journal of Big Data. – 2023. – URL: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00796-3 (дата обращения: 21.04.2025).

10. Botometer by OSoMe // Indiana University. – URL: https://botometer.osome.iu.edu/ (дата обращения: 21.04.2025).

References

1. Rozanova, A. V. Fraudsters began to steal money through a fake Russian Post bot. *RBC Life*, 2025. Available at: https://www.rbc.ru/life/news/67c7ef979a79472746515cc2 (accessed 20.04.2025) (In Russ.).

2. Borodovsky, D. A. In 2024, fraudsters stole 168 billion rubles from Russians. *Hi-Tech Mail*, 2025. Available at: https://hi-tech.mail.ru/news/120371-za-2024-god-moshenniki-ukrali-u-rossiyan-168-mlrd-rublej/ (accessed 20.04.2025) (In Russ.).

3. Anisimova, N. A. Data of accounts spreading fakes about coronavirus were transferred to security forces. *RBC.RU*, 2020. Available at: https://www.rbc.ru/rbcfreenews/5e5fbff99a7947c60350de28 (accessed 20.04.2025) (In Russ.).

4. Kaminsky, S. N. Fraud with the Russian Post bot in Telegram. *kaspersky daily*, 2025. Available at: https://www.kaspersky.ru/blog/russian-post-telegram-scam/39383/ (accessed 20.04.2025) (In Russ.).

5. Samokhvalov, D. A. Machine Learning-Based malicious users detection in the VKontakte social network. *Semantic Scholar*, 2020. Available at: https://pdfs.semanticscholar.org/5db8/3d3e88b653e1d2242d547420d99f06a09228.pdf (accessed 20.04.2025) (In Russ.).

6. Chastikova, V. A., Zherlitsyn, S. A., Mityugov, S. A. *Artificial intelligence technologies in information security : monograph.* Krasnodar, Publishing house of KubSTU, 2024. 315 p. (In Russ.).

7. Chechulin, A. A., Kolomeets, M. V. Approach to detecting malicious bots in the social network Vkontakte and assessing their parameters. *Proceedings of educational institutions of communication*, 2024, vol. 10, no. 2, pp. 92–101. DOI: 10.31854/1813-324X-2024-10-2-92-101. Available at: https://www.researchgate.net/publication/380418227 (accessed 20.04.2025) (In Russ.).

8. Ng, L., Carley, K. What is a Social Media Bot? A Global Comparison of Bot and Human Characteristics. *arxiv.org*, 2025. Available at: https://arxiv.org/html/2501.00855v1 (accessed 20.04.2025).

9. Denghan, A., Suita, K. Detecting bots in social-networks using node and structural embeddings. *Journal of Big Data*, 2023. Available at: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00796-3 (accessed 21.04.2025).

10. Botometer by OSoMe. *Indiana University*. Available at: https://botometer.osome.iu.edu/ (accessed 21.04.2025).

Статья поступила в редакцию 27.05.2025; одобрена после рецензирования 25.07.2025; принята к публикации 04.08.2025.

The article was submitted 27.05.2025; approved after reviewing 25.07.2025; accepted for publication 04.08.2025.

УДК 004.056

РЕШЕНИЕ ЗАДАЧИ ОПТИМИЗАЦИИ ПОДБОРА МЕР ДЛЯ ПОВЫШЕНИЯ ОЦЕНКИ СООТВЕТСТВИЯ ТРЕБОВАНИЯМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ КРЕДИТНЫХ ОРГАНИЗАЦИЙ

Жукова Марина Николаевна, Сибирский государственный университет науки и технологий, 660037, Российская Федерация, г. Красноярск, пр. им. газ. «Красноярский рабочий», 31,
кандидат технических наук, доцент, ORCID: 0000-0003-3441-3041, e-mail: zhukova@sibsau.ru
Яковлева Анастасия Олеговна, Сибирский государственный университет науки и технологий, 660037, Российская Федерация, г. Красноярск, пр. им. газ. «Красноярский рабочий», 31,
аспирант, ORCID: 0009-0008-7828-0977, e-mail: petrova754@mail.ru

В работе рассматривается задача оптимизации подбора мер защиты путем построения и применения аналитического модуля в составе инструмента автоматизации работы со стандартами серии ГОСТ Р 57580.1 и 57580.2. Модуль реализует функционал по приведению в соответствие требованиям информационной безопасности инфраструктуры кредитной организации в случае получения неудовлетворительных результатов оценки соответствия. Приводится описание критериев, допустимых к использованию при формировании сценариев повышения оценки, которые возможно применить в отсутствии возможности получения полной информации о системе защиты, основываясь только на результатах проведенного в организации аудита. Выбранный подход по получению сценариев доработки, который представляет собой создание наборов мер защиты, предлагаемых для выбора и реализации в соответствии с уровнем защищенности, основан на анализе зависимостей и взаимосвязей между параметрами оценки. В статье показано поэтапное описание и порядок работы компонентов аналитического модуля, реализующего предложенный алгоритм формирования сценариев для приведения инфраструктуры кредитной организации в соответствие требованиям стандартов.

Ключевые слова: оценка соответствия, кредитные организации, аналитический модуль, корреляционные коэффициенты, уровень защищенности

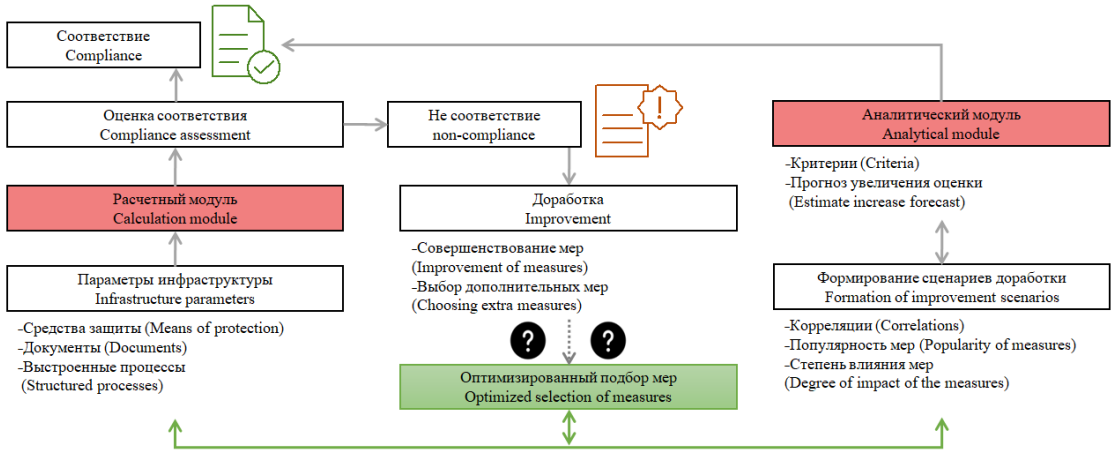
SOLVING THE PROBLEM OF OPTIMIZING MERDLING SELECTION TO IMPROVE COMPLIANCE WITH INFORMATION SECURITY REQUIREMENTS OF CREDIT INSTITUTIONS

Zhukova Marina N., Reshetnev Siberian State University of Science and Technology, 31 Krasnoyarsky Rabochy Ave., Krasnoyarsk, 660037, Russian Federation,
Cand. Sci. (Engineering), Associate Professor, ORCID: 0000-0003-3441-3041, e-mail: zhukova@sibsau.ru
Iakovleva Anastasia O., Reshetnev Siberian State University of Science and Technology, 31 Krasnoyarsky Rabochy Ave., Krasnoyarsk, 660037, Russian Federation,
graduate student, ORCID: 0009-0008-7828-0977, e-mail: petrova754@mail.ru

The paper considers the problem of optimizing the selection of protection measures by building and applying an analytical module as part of an automation tool for working with GOST R 57580.1 and 57580.2 series standards. The module implements the functionality to bring the information security requirements of the credit institution's infrastructure in case of unsatisfactory compliance assessment results. The criteria acceptable for use in the formation of evaluation enhancement scenarios are described, which can be applied in the absence of the possibility of obtaining complete information about the protection system based only on the results of an audit conducted in the organization. The chosen approach to obtaining refinement scenarios, which is the creation of sets of protection measures proposed for selection and implementation in accordance with the security level, is based on an analysis of the dependencies and relationships between the evaluation parameters. The article shows a step-by-step description and operation of the components of the analytical module, which implements the proposed algorithm for generating scenarios to bring the infrastructure of a credit institution in line with the requirements of the standards.

Keywords: conformity assessment, financial institution, analytics module, correlation coefficients, security level

Graphical annotation (Графическая аннотация)



ВВЕДЕНИЕ

Вопросам проведения процедур оценки соответствия требованиям к защите информации уделяется большое внимание. Интерес исследователей привлекают как общие подходы и методики, позволяющие оценить текущее состояние безопасности информационной системы предприятия на предмет соответствия требованиям нормативных и руководящих документов [1], так и аудит определенных направлений информационных систем [2, 3].

Для кредитных организаций выполнение требований стандартов и поддержание требуемого уровня защищенности систем и сервисов является обязательной и приоритетной задачей, поскольку напрямую влияет на наличие лицензии для осуществления деятельности или отдельных видов операций и возможность их дальнейшей работы. Описание проведения оценки соответствия приводится в стандарте ГОСТ Р 57580.2 – 2018 «Безопасность финансовых (банковских) операций. Защита информации финансовых организаций. Методика оценки соответствия» [4], а информация о требованиях к защите в ГОСТ Р 57580.1 – 2017 «Безопасность финансовых (банковских) операций. Защита информации финансовых организаций. Базовый набор организационных и технических мер» [5].

При неудовлетворительных результатах оценки перед специалистами возникает задача ее повышения [6]. Разработано программное решение [7], которое позволяет рассчитать соответствие системы безопасности кредитной организации требуемому уровню защищенности по каждому из восьми оцениваемых процессов. Решение реализует расчетную часть инструмента в рамках общего подхода работы со стандартами [8], который предполагает автоматизацию процедуры проведения оценки соответствия и, помимо расчетного модуля, предусматривает наличие модуля аналитики.

При практическом применении стандарта и автоматизированных решений основной проблемой является не процесс расчета оценки, а поиск оптимальных стратегий ее повышения в случае получения недостаточного показателя соответствия. Модуль аналитики необходим для формирования сценариев повышения оценки соответствия, возможных к применению в инфраструктуре кредитной организации, на основе результатов расчета текущей оценки, для достижения требуемого уровня соответствия требованиям информационной безопасности.

Таким образом, целью исследования является создание и описание алгоритма работы проектируемого аналитического модуля, применение которого в совокупности с результатами модуля расчета позволит автоматизировать процесс работы с задачей соответствия требованиям стандартов ГОСТ Р 57580.1 и ГОСТ Р 57580.2.

Разработка такого компонента невозможна без описания критериев выбора мер защиты, которое было представлено ранее [9], и описания степени влияния выполнения каждой меры на итоговый результат оценки [10]. В исследовании будет произведено уточнение критериев, описание выбора дополнительных параметров и их упорядочивание для создания и проектирования аналитического модуля.

ОБЩЕЕ ОПИСАНИЕ ИНСТРУМЕНТА АВТОМАТИЗАЦИИ. ПОЗИЦИЯ АНАЛИТИЧЕСКОГО МОДУЛЯ

Решение по автоматизации процесса работы с инфраструктурой в соответствии с требованиями серии стандартов 57580.X представлено на рисунке 1 и включает в себя два основных модуля:

1) расчетный модуль (калькулятор) для оценки ситуации по выполнению требований;
2) аналитический модуль подбора дополнительных (корректирующих до конкретного уровня) мер для получения нужного показателя оценки соответствия [8].

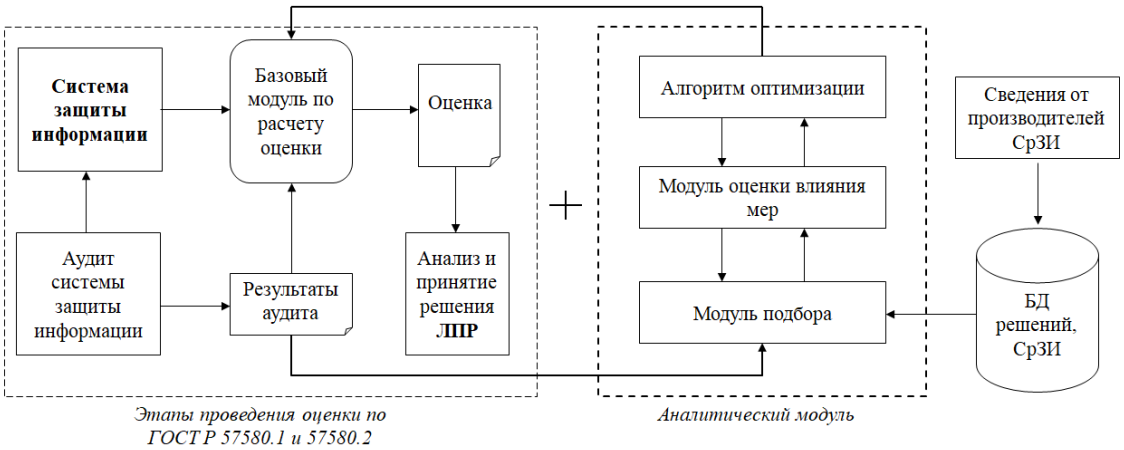


Рисунок 1 – Общая схема решения

Работу предлагаемого решения, представленного на рисунке выше, можно описать следующим образом:

1. Сведения об имеющейся системе защиты финансовой организации анализируются специалистами на этапе предварительного аудита и загружаются в базовый модуль расчета, который представляет собой «калькулятор», позволяющий вычислить оценку для контура исходя из выбранного уровня защиты на текущий момент.

2. По результатам работы расчетного модуля выводятся результаты аудита, которые содержат информацию о параметрах, имеющих низкий показатель соответствия (потенциальные места доработки), в форме таблицы, которую можно сразу отдать лицу, принимающему решение (ЛПР), и на этом закончить работу, а можно загрузить в аналитический модуль;

3. В **аналитическом модуле** присутствуют:

– модуль оценки мер, где хранятся рассчитанные коэффициенты, отражающие степень влияния меры на результат оценки при ее реализации;

– модуль, отвечающий за оптимизацию, который содержит информацию о зависимостях оцениваемых параметров и на основе этой информации предлагает меры, наиболее подходящие для доработки или реализации;

– модуль подбора, в котором содержится информация о целевом алгоритме (функции) подбора мер защиты – критериях формирования сценариев доработки;

– дополнительно: перечень выбранных мер для доработки можно сравнить со списком решений, предлагаемых производителями, которые закрывают те или иные требования (однако на текущий момент готовой базы таких решений нет), предложить список средств для реализации каждой меры и результаты этой работы отдать ЛПР, либо загрузить полученную информацию в аналитический модуль в перспективе в качестве дополнительного параметра.

АНАЛИТИЧЕСКИЙ МОДУЛЬ. ОПИСАНИЕ КОМПОНЕНТОВ

Аналитический модуль, исходя из цели и задачи его разработки, представляет собой инструмент автоматизации, предназначенный для того, чтобы формировать наборы мер, в дальнейшем именуемые как сценарии, предназначенные для применения в инфраструктуре для повышения оценки соответствия. Согласно рисунку 1, в модуле представлены три основных компонента: модуль оценки мер, модуль оптимизации, модуль подбора, которые требуют подробного описания.

Модуль оценки мер предназначен для хранения информации о весовых коэффициентах, отражающих степень влияния выполнения или реализации меры на итоговый показатель оценки. При подборе мер для сценария повышения оценки потребуется понимание того, насколько возрастет значение оценки после выполнения предложенных мер, из чего можно будет сделать вывод о достаточности предлагаемого набора. Подробный порядок расчета и итоговые показатели весовых коэффициентов для каждой меры были отдельно опубликованы [10].

Модуль оптимизации предполагает хранение параметров мер, исходя из которых принимается решение об их включении в сценарий доработки. Таких параметров было выделено три: популярность меры к реализации и наличие зависимостей между оцениваемыми параметрами, которое подразделяется еще на два блока: зависимость от применяемых средств защиты и зависимость от выполнения других мер. Предположение о возможности существования таких зависимостей высказывалось ранее [11], однако подробного описания таких зависимостей не приводилось.

Идея поиска зависимостей между оцениваемыми параметрами (мерами защиты) в том, чтобы использовать информацию, полученную на этапе расчета соответствия и на ее основе предложить

меры для доработки или реализации. Требуется изучить следующие предположения по анализу зависимостей и взаимосвязей между мерами:

1) применения имеющегося набора средств защиты для закрытия других мер, на основании предположения о том, что решения по защите информации способны одновременно справляться с несколькими задачами, что позволяет применять их для комплекса различных мер, причем сразу в нескольких процессах;

2) доработки мер или выбора часто реализуемых мер, которые были оценены как «не выбранные» (на 0) или в рамках процессного подхода были оценены как «не реализованные» или «частично реализованные» (на 0 или 0,5), на основании вероятностного подхода из экспертных оценок;

3) выбора и реализации мер, связанных с теми, что уже включены в оценку на основании анализа зависимостей между ними, который основывается на предположении, что меры защиты связаны между собой, поскольку создание и построение систем защиты выполняется комплексно, а значит, реализация одних мер косвенно ведет к реализации других хотя бы в рамках одного процесса [11].

Для поиска зависимостей между оцениваемыми параметрами потребуется разметка данных, которую можно выполнить вручную или с применением средств автоматизации. В настоящем исследовании был применен гибридный метод обработки данных: сначала произведен автоматический поиск зависимостей, а затем найденные зависимости изучены авторами вручную для поиска ошибочно найденных зависимостей и их устранения.

Данные для анализа были получены в результате обратной связи, полученной от специалистов банков при тестировании модуля расчета. Полученная база знаний использовалась для тестового режима и проектирования модуля аналитики. Предполагается, что использование инструмента кредитными организациями на постоянной основе позволит расширять базу и совершенствовать разрабатываемый инструмент, в особенности его аналитический модуль.

Этап 1. Анализ популярности меры. Анализ популярности представлен на примере четырех оценок для демонстрации логики его проведения. Был выполнен расчет популярности для выполнения каждой меры, при этом мера считалась непопулярной для выполнения и реализации только в случае, когда была оценена 0, что показано в таблице 1 с примерами результатов расчета (см. предположение 2 о популярности реализации определенных мер защиты или их потенциале к доработке [11]).

Таблица 1 – Макет таблицы расчета популярности мер (на примере процесса 8)

Меры защиты оценки	**УЗП 1**	**…**	**ИУ 8**	**ПЗИ 1**	**…**	**СЗИ 4**
Оценка 1	1			0		1
Оценка 2	1			0,5		0,5
Оценка 3	0			1		1
Оценка 4	1			1		1
Вероятность рекомендации меры (в %)	0,75			0,75		1

Анализ зависимостей был произведен путем расчета корреляции по методу Пирсона между мерами, с применением метода df.corr библиотеки pandas. Альтернативные расчеты для контроля точности анализа были произведены путем использования метода корреляционного анализа в инструменте Loginom. Размерность полученных карт и таблиц большая, что не позволяет полностью вместить их в описание работы.

Этап 2. Поиск зависимостей в рамках каждого процесса. Для поиска зависимостей внутри процесса учитывались зависимости как между мерами защиты $E_{МЗИ}$, так и между мерами $E_{МОУ}$ [4, 5], отвечающими за процессный подход к защите (исключение: мера – к самой себе). В рамках одного процесса такая зависимость предполагалась на основании комплексного подхода к созданию систем защиты (см. предположение 3 о возможности доработки мер и зависимости между мерами внутри одного процесса [11]).

Реализован анализ в виде таблиц, пример строки такой таблицы приведен ниже (табл. 2). Столбцы этих таблиц содержат те же значения, что в строках, но транспонированные.

Таблица 2 – Макет строки таблицы для тепловой карты / карты корреляционных коэффициентов (на примере процесса 8)

Процесс 8 «Защита информации при осуществлении удаленного логического доступа с использованием мобильных (переносных) устройств»														
ЗУД 1	…	ЗУД 12	ПЗИ 1	…	ПЗИ 5	РЗИ 1	…	РЗИ 16	КЗИ 1	…	КЗИ 12	СЗИ 1	…	СЗИ 4

Этап 3. Поиск зависимостей в между процессами. Анализ зависимостей между процессами производился без учета мер процессного подхода (исключение: мера – к самой себе), поскольку предполагать зависимости между планированием, реализацией, контролем и совершенствованием мер защиты, которые в основе своей реализуются организационно, оснований не было (см. предположение 1 о зависимости по применяемым средствам защиты [11]). Реализован анализ в виде таблиц, пример строки такой таблицы приведен ниже (табл. 3). Столбцы этих таблиц содержат те же значения, но транспонированные.

Таблица 3 – Макет строки таблицы для тепловой карты / карты корреляционных коэффициентов для поиска межпроцессных зависимостей между мерами

Процесс 1			**Процесс 2**			**…**	**Процесс 7**			**Процесс 8**		
УЗП 1	…	ИУ 8	СМЭ 1	…	ЗБС 10	…	ЗСВ 1	…	ЗСВ 43	ЗУД 1	…	ЗУД 12

Для примера полученных результатов анализа зависимостей, полученных на этапах 2 и 3, на рисунке 2 приводится фрагмент тепловой карты для процесса 6 «Управление инцидентами защиты информации». Также на рисунке приведена шкала ослабления зависимостей мер друг от друга (от сильных зависимостей, близких к «1» – красный цвет, до слабых и обратных, близких к «–1» – синий цвет). В настоящем исследовании найденные отрицательные зависимости не исследовались, поскольку наличие таких зависимостей предполагает, что выполнение одной меры защиты отрицательно влияет на выполнение других мер защиты, что противоречит логике стандартов ГОСТ Р 57580.1 и ГОСТ Р 57580.2.

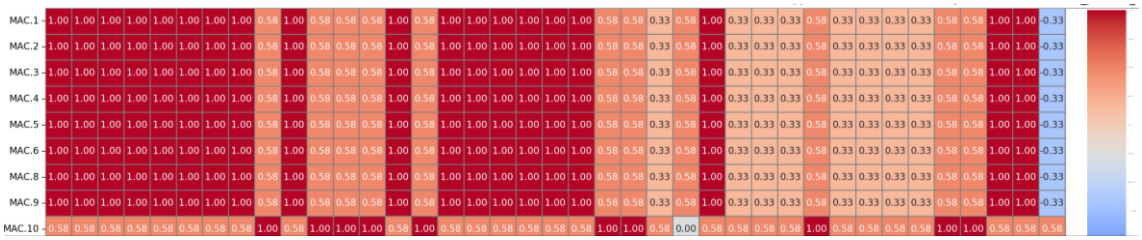


Рисунок 2 – Фрагмент тепловой карты на примере процесса 6

Критерием рассмотрения результата являются переменные с положительными зависимостями, где «коэффициент корреляции Пирсона (ρ) > 0,7», поскольку такое значение коэффициента свидетельствует о наличии «заметной корреляции» [12].

Зависимости с подходящим коэффициентом анализировались вручную: вычитывались на предмет возможности существования такой связи с точки зрения логического объяснения. Большинство из них подтвердились. Результаты расчетов, полученные в Loginom и с применением df.corr из pandas, не противоречат друг другу и, после детального анализа коэффициентов корреляции для каждого процесса в отдельности и для всех процессов в совокупности, были добавлены в модуль подбора как справочные таблицы.

Модуль подбора в отличие от первых двух, которые являются информационными, содержит непосредственно целевой алгоритм, функцию формирования сценариев повышения оценки соответствия и критериев добавления в них мер защиты.

При изначальном проведении анализа отмечалось два критерия, которые можно применить в отсутствии информации об оцениваемой инфраструктуре:

1) скорость реализации, с учетом приблизительных подсчетов времени для реализации различных мер защиты со стороны экспертного сообщества;

2) подбор планируемых изменений, который можно реализовать в нескольких вариациях:

– путем поиска всех возможных комбинаций мер, которые в сумме дадут недостающую часть в оценке;

– по оценке достижения результата за счет конкретного выбора категории мер (организационные или технические);

– по выбору количества параметров, подвергаемых изменению [9].

Несмотря на то, что уже существуют приблизительные подсчеты времени, необходимые для реализации различных мер защиты со стороны экспертного сообщества [13, 14], пересчитать эти показатели для специалистов конкретной организации и инфраструктуры проблематично, а значит, критерий скорости реализации неприменим.

Таким образом, остается один критерий из предложенных, который можно использовать, – «Подбор планируемых изменений». Однако потребуется его корректировка, чтобы в основу заложить корреляционные зависимости оцениваемых параметров, при этом изначальная основа по оценке веса мер будет применена. Критерий также позволит формировать комбинации – сценарии приведения оценки к требуемому уровню, а возможные варианты сбора этих комбинаций будут основаны на следующих основных параметрах, которые представляют собой целевую функцию (алгоритм), которая реализуется в модуле подбора и выглядит следующим образом:

1. Приоритет для предложения мер при формировании сценария повышения оценки должен быть выставлен в зависимости от коэффициента корреляции меры для добавления ее в сценарий.
2. Добавление мер в сценарий должно производиться согласно корреляционному коэффициенту (от большего к меньшему), с учетом порогового значения, тогда оптимально подходящие меры для доработки и реализации будут предложены в первую очередь.
3. На каждом этапе добавления мер в сценарий должен производиться подсчет увеличения показателя оценки за счет применения информации об их весовых коэффициентах [10].
4. Коррелирующие меры добавляются в сценарий повышения оценки до тех пор, пока не будет получен требуемый результат (достигнута разница между требуемым значением оценки и текущим ее показателем).
5. Если коррелирующих мер не хватает для получения требуемой оценки, следует предлагать «популярные» меры, полученные при анализе оценок для каждого процесса.
6. А если и этих мер будет недостаточно для достижения требуемого показателя соответствия, пользователю предлагаются меры, имеющие больший весовой коэффициент в оценке, которые можно добавить в сценарий повышения оценки после коррелирующих и «популярных».

ПРИМЕНЕНИЕ КОМПОНЕНТОВ В РАЗРАБОТКЕ АНАЛИТИЧЕСКОГО МОДУЛЯ

Аналитический модуль необходим для решения задачи приведения оценки соответствия к требуемому уровню, с опорой на параметры инфраструктуры, получаемые из текущего (недостающего) показателя оценки. Схема модуля представлена на рисунке 3, а алгоритм его работы следующий:

1. В модуль загружаются результаты проведения внутреннего аудита, полученные в результате работы расчетного модуля;
2. На основе оценки из расчетного модуля и требуемого результата (уровня) оценки соответствия для каждого процесса вычисляется числовое значение разницы текущего уровня оценки и требуемого по каждому процессу в рамках GAP-анализа.
3. На основе весов мер, зависимостей, полученных путем корреляционного анализа и анализа «популярных» мер, подбираются меры, предлагаемые к доработке или реализации.
4. Подобранные меры (все возможные) в форме перечня выводятся пользователю в читаемой и заранее определенной форме с указанием прироста оценки при выполнении каждой меры, рассчитанного исходя из весовых коэффициентов; затем пользователю предоставляется выбор: самостоятельно анализировать список и выбирать подходящие меры или же попросить модуль сформировать «сценарий» доработки, где будет представлена комбинация мер, собранная и ранжированная по определенным в модуле подбора критериям (целевой функции), которые в сумме дадут недостающую часть оценки.
5. Сформированный сценарий выводится как набор мер, рекомендуемых к доработке в порядке приоритета их реализации, согласно критериям алгоритмической целевой функции.

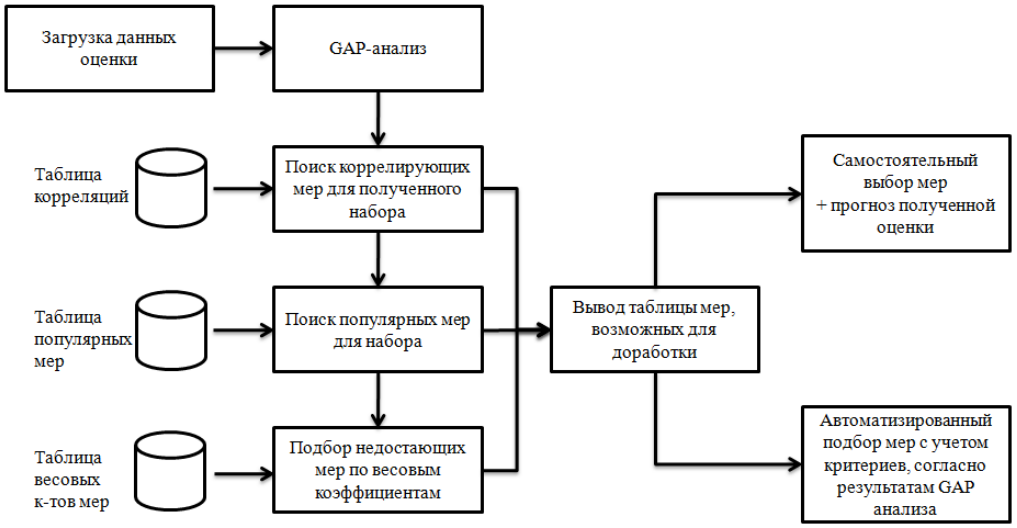


Рисунок 3 – Этапы работы аналитического модуля

На этапах представленного алгоритма возникают задачи, которые требуется решить.

1 – Получение параметров оценки.

На этапе 1 производится загрузка результатов оценки из расчетного модуля, включая итоговый показатель соответствия, и справочной таблицы с описанием мер защиты в определенном формате.

2 – GAP-анализ.

На этапе 2 выполняется операция вычитания (нахождения разности) между показателем уровня оценки (требуемого) и уровня оценки (текущего), чтобы получить информацию о недостающем коэффициенте оценки.

3 – Получение параметров мер.

На этапе 3 загружаются данные из информационных модулей в модуль подбора (для статьи используется пример одного из уровней защиты – 2), среди которых:

– корреляционные коэффициенты, отражающие зависимости мер (в рамках одного процесса и между процессами), пример таблицы для которых приведен на рисунке 4;

– корреляционные (вероятностные) коэффициенты, отражающие популярность мер, пример таблицы для которых приведен на рисунке 5;

– коэффициенты «вклада» меры в доработку (степень влияния меры), расчет коэффициентов для которых был приведен [10] и в каждой таблице выполняется динамически в зависимости от количества применимых к инфраструктуре мер, пример таблицы приведен на рисунке 6.

Уровень защиты информации - 2										
Мера 1					Мера 2					
Код процесса	Условное обозначение и номер меры	Содержание мер системы защиты информации	Способ реализации (О/Т)	Средство защиты / орг документ	Код процесса	Условное обозначение и номер меры	Содержание мер системы защиты информации	Способ реализации (О/Т)	Средство защиты / орг документ	Корреляция
6	МАС 1	Организация мониторинга данных регистрации о событиях защиты информации, формируемых техническими мерами, входящими состав систем защиты информации	Т	SIEM	6	РИ 5	Установление и применение единых правил регистрации и классификации инцидентов защиты информации в части состава и содержания атрибутов,описывающих инцидент защиты информации и их возможных значений	Т	SIEM	1

Рисунок 4 – Пример заполненной строки таблицы корреляции

На рисунке 4, который отражает пример строки заполненной таблицы, присутствуют столбцы, отмеченные серым цветом. Эти столбцы разработаны с учетом перспективы появления таблиц соответствия мер защиты и средств или организационных документов, которые их закрывают. На текущий момент такой таблицы нет, но в процессе использования инструмента автоматизации ее можно будет сформировать и заполнить.

Уровень защиты информации - 2				
Мера 1				
Код процесса	Условное обозначение и номер меры	Содержание мер системы защиты информации	Способ реализации (О/Т)	Популярность коэффициент
6	МАС 1	Организация мониторинга данных регистрации о событиях защиты информации, формируемых техническими мерами, входящими состав систем защиты информации	Т	0, 87

Рисунок 5 – Пример заполненной строки таблицы популярных мер

Уровень защиты информации - 2				
Мера				Весовой коэффициент
Код процесса	Условное обозначение и номер меры	Содержание мер системы защиты информации	Способ реализации (О/Т)	
6	МАС 1	Организация мониторинга данных регистрации о событиях защиты информации, формируемых техническими мерами, входящими состав систем защиты информации	Т	0,5* ($E_{МЗИ_j}$/ N) Рассчитывается динамически [10]

Рисунок 6 – Пример заполненной строки таблицы весов

4 – Формирование сценариев.

На этапе 4 в рамках решения задачи повышения оценки для каждого процесса пользователю выводится перечень мер внутри процесса, для доработки или реализации. Затем, согласно целевой функции (алгоритму), описанной в модуле подбора, формируется сценарий доработки, позволяющий получить требуемый показатель оценки уровня соответствия. Для того чтобы утверждать, что мер из сценария будет достаточно, требуется выполнение условия, что показатель прироста должен быть больше или равен недостающему коэффициенту, полученному на этапе 2. Пользователь наводит указатель на ячейку предлагаемой для доработки или реализации меры, что показано на рисунке 7, на котором представлена следующая информация:

– обозначение меры;
– описание меры;
– требование к реализации меры (О/Т);
– текущий показатель оценки: выбор и реализация самой меры (0/1 или 0/0,5/1);
– планируемый показатель оценки: выбор и реализация самой меры (0/1 или 0/0,5/1);
– предполагаемый прирост оценки по тем же коэффициентам (при увеличении показателя по оценке и, как следствие, увеличении степени прироста показателя по весу), рассчитанный динамически для каждой меры.

Уровень защиты информации - 2					
Мера 1					
Условное обозначение и номер меры	Содержание мер системы защиты информации	Способ реализации (О/Т)	Текущий показатель оценки	Планируемый показатель оценки	Прирост оценки
ПЗИ 1	Документарное определение области применения процесса системы защиты информации для уровней информационной инфраструктуры, определенных в 6.2 настоящего стандарта	О	0	1	0,01

Рисунок 7 – Пример вывода таблицы мер

5 – Принятие решения.

На основе информации этапа 4 пользователь сможет решить, какие из мер он будет реализовывать или дорабатывать для получения нужной оценки. Возможны два пути:

– воспользоваться готовым, предложенным аналитическим модулем, сценарием получения оценки требуемого уровня, который будет выглядеть как пример, приведенный на рисунке 7;
– самостоятельно «отметить» понравившиеся меры (из перечня всех доступных мер для доработки) и, после выбора этих мер, по нажатию кнопки, получить итоговый показатель прироста оценки, что позволит оценить достаточность (полноту) выбранного самостоятельно набора (рис. 8).

Уровень защиты информации - 2						
Мера 1						
Условное обозначение и номер меры	Содержание мер системы защиты информации	Способ реализации (О/Т)	Текущий показатель оценки	Планируемый показатель оценки	Прирост оценки	Добавить меру?
ПЗИ 1	Документарное определение области применения процесса системы защиты информации для уровней информационной инфраструктуры, определенных в 6.2 настоящего стандарта	О	0	1	0,01	☐

Рисунок 8 – Самостоятельный выбор мер доработки

Если у пользователя недостаточно времени, желания, квалификации для самостоятельного формирования «набора доработок», аналитический модуль предлагает готовый оптимальный и эффективный сценарий повышения оценки, полученный в результате решения задачи многокритериального анализа данных.

ЗАКЛЮЧЕНИЕ

Проблема получения низких результатов при проведении аудита на соответствие требованиям стандартов является актуальной для специалистов информационной безопасности кредитных организаций. Разработанный аналитический модуль позволяет получить оптимизированный набор мер защиты, возможных для выбора или реализации в инфраструктуре, основываясь на ее параметрах.

В работе произведен выбор критерия формирования сценариев повышения оценки, для применения которого был выполнен анализ зависимостей между мерами защиты и анализ популярности мер для их выполнения. Также применяются предшествующие результаты исследований авторов по расчету степени влияния выполнения мер защиты на итоговый показатель соответствия. Описана целевая функция, алгоритм получения наборов мер защиты, предлагаемых для повышения оценки.

Применение алгоритма позволяет оптимизировано предлагать меры для доработки, исходя из их взаимосвязи с уже реализованными и оцененными параметрами системы защиты организации.

Возможность проведения аудита с применением модуля расчета и последующего формирования сценариев повышения оценки соответствия с применением разработанного в настоящей статье аналитического модуля позволяет пользователю инструмента автоматизировать процесс работы с задачей соответствия требованиям стандартов ГОСТ Р 57580.1 и ГОСТ Р 57580.2, что свидетельствует о достижении цели настоящей статьи.

Список источников

1. Марков, О. Н. Оценка соответствия информационной безопасности объектов аудита требованиям нормативных документов / О. Н. Марков, Е. В. Кусов, Я. М. Гвоздик // Проблемы информационной безопасности. Компьютерные системы. – 2006. – № 3. – С. 62–71.

2. Баранкова, И. И. Аудит информационной безопасности промышленных предприятий, направленный на оценку соответствия требованиям российского и международного законодательства / И. И. Баранкова, Е. А. Семавина, У. В. Михайлова // Вестник УрФО. Безопасность в информационной сфере. 2022. № 3 (45). С. 76–82.

3. Попов, И. Ю. Разработка метода оценки соответствия обеспечения безопасности персональных данных в информационных системах согласно требованиям регулятора / И. Ю. Попов // Сборник трудов V Всероссийского конгресса молодых ученых. – 2016. – Т. 2. – С. 94–98.

4. ГОСТ Р 57580.2 – 2018 (дата введения 28.03.2018). – URL: https://rosgosts.ru/file/gost/03/060 /gost_r_57580.2-2018.pdf (дата обращения: 24.05.2025).

5. ГОСТ Р 57580.1 – 2017 (дата введения 08.08.2017). – URL: https://docs.cntd.ru/document/1200146534 (дата обращения: 24.05.2025).

6. Яковлева, А. О. О задаче поиска оптимальных стратегий для повышения оценки соответствия требованиям информационной безопасности для кредитных организаций / А. О. Яковлева // Актуальные проблемы авиации и космонавтики : сборник материалов XI Международной научно-практической конференции (в печати).

7. Яковлева, А. О. Разработка программного решения для формирования адаптированного набора мер при проведении оценки соответствия по ГОСТ Р 57580.2 в зависимости от требований ГОСТ Р 57580.1 / А. О. Яковлева, М. Н. Жукова // Современные методы, средства и технологии защиты информации – 2024 : материалы XV Международной научно-практической конференции имени Олега Борисовича Макаревича. – URL: https://ictis.sfedu.ru/wp-content/uploads/2024/12/XV_Конференция_ИБ-2024.pdf(дата обращения: 03.06.2025).

8. Яковлева, А. О. Разработка подхода к анализу соответствия результатов проведения аудита и применяемых оценочных процедур в рамках GAP анализа для кредитно-финансовой сферы / А. О. Яковлева, М. Н. Жукова // Вестник УрФО. Безопасность в информационной сфере. 2024. № 3 (53). – URL: https://info-secur.ru/index.php/ojs/article/view/476 (дата обращения: 24.06.2025).

9. Iakovleva, A. On the general approach and criteria for optimizing the selection of protective measures in preparation for the procedure for assessing compliance with information security requirements for financial organizations / A. Iakovleva, M. Zhukova // AISMA-2025: International Workshop on Advanced Information Security Management and Applications (in print).

10. Яковлева, А. О. Расчет коэффициентов, отражающих степень влияния выполнения мер защиты информации при проведении оценки соответствия для финансовых организаций / А. О. Яковлева, М. Н. Жукова // Прикаспийский журнал: управление и высокие технологии. – 2024. – № 4 (68). – С. 29–35.

11. Яковлева, А. О. О вопросе применения методов корреляционного анализа для работы с параметрами оценки соответствия требованиям информационной безопасности кредитных организаций / А. О. Яковлева // Актуальные проблемы авиации и космонавтики : сборник материалов XI Международной научно-практической конференции (в печати).

12. Расчет корреляции в Pandas. – URL: https://sky.pro/wiki/analytics/raschet-korrelyatsii-v-pandas-metod-dfcorr()-dlya-analiza-dannyh/ (дата обращения: 15.07.2025).

13. УЦСБ. Серия вебинаров «Безопасность финансовых организаций». – URL: https://www.ussc.ru /events/zapisi-vebinarov/seriya-vebinarov-bezopasnost-finansovykh-organizatsiy/ (дата обращения: 20.07.2025).

14. Как повысить оценку ГОСТ 57580: быстро, эффективно, с первого раза. – URL: https://bi.zone/expertise/events/kak-povysit-otsenku-gost-57580-bystro-effektivno-s-pervogo-raza/ (дата обращения: 23.07.2025).

References

1. Markov, O. N., Kusov, E. V., Gvozdik, Ya. M. Assessment of compliance of information security of audit objects with the requirements of regulatory documents. *Problems of Information Security. Computer Systems*, 2006, no. 3, pp. 62–71 (In Russ.).

2. Barankova, I. I., Semavina, E. A., Mikhailova, U. V. Information Security Audit of Industrial Enterprises Aimed at Assessing Compliance with Russian and International Legislation. *Bulletin of the Ural Federal District. Information Security*, 2022, no. 3 (45), pp. 76–82 (In Russ.).

3. Popov, I. Yu. Development of a method for assessing compliance of personal data security in information systems with regulator requirements. *Collection of works of the V All-Russian congress of young scientists*, 2016, vol. 2, pp. 94–98 (In Russ.).

4. *GOST R 57580.2 – 2018* (date of introduction: 28.03.2018). Available at: https://rosgosts.ru/file/gost/03/060 /gost_r_57580.2-2018.pdf (accessed 24.05.2025) (In Russ.).

5. *GOST R 57580.1 – 2017* (date of introduction: 08.08.2017). Available at: https://docs.cntd.ru/document/1200146534 (accessed 24.05.2025) (In Russ.).

6. Yakovleva, A. O. On the problem of searching for optimal strategies to improve the assessment of compliance with information security requirements for credit institutions. *Actual problems of aviation and cosmonautics: collection of materials from the XI International scientific and practical conference* (in press) (In Russ.).

7. Yakovleva, A. O., Zhukova, M. N. Development of a software solution for generating an adapted set of measures for conducting conformity assessment according to GOST R 57580.2 depending on the requirements of GOST R 57580.1. Modern methods, means and technologies of information security – 2024 : proceedings of the XV International scientific and practical conference named after Oleg Borisovich Makarevich. Available at: https://ictis.sfedu.ru/wp-content/uploads/2024/12/XV_Конференция_ИБ-2024.pdf (accessed 03.06.2025) (In Russ.).

8. Yakovleva, A. O., Zhukova, M. N. Development of an approach to the analysis of compliance of audit results and applied assessment procedures within the framework of GAP analysis for the credit and financial sphere. *Bulletin of the Ural Federal District. Security in the information sphere*, 2024, no. 3 (53). Available at: https://info-secur.ru/index.php/ojs/article/view/476 (accessed 24.06.2025) (In Russ.).

9. Iakovleva, A., Zhukova, M. On the general approach and criteria for optimizing the selection of protective measures in preparation for the procedure for assessing compliance with information security requirements for financial organizations. *AISMA-2025: International Workshop on Advanced Information Security Management and Applications* (in print).

10. Yakovleva, A. O. Zhukova, M. N. Calculation of coefficients reflecting the degree of influence of the implementation of information security measures during the assessment of compliance for financial organizations. *Caspian Journal: Control and High Technologies*, 2024, no. 4 (68), pp. 29–35 (In Russ.).

11. Yakovleva, A. O. On the issue of applying correlation analysis methods to work with parameters for assessing compliance with information security requirements of credit institutions. *Actual problems of aviation and cosmonautics: collection of materials from the XI International scientific and practical conference* (in press) (In Russ.).

12. *Correlation calculation in Pandas*. Available at: https://sky.pro/wiki/analytics/raschet-korrelyatsii-v-pandas-metod-dfcorr()-dlya-analiza-dannyh/ (accessed 15.07.2025) (In Russ.).

13. *UCSB. Webinar series "Security of financial organizations"*. Available at: https://www.ussc.ru/events/zapisi-vebinarov/seriya-vebinarov-bezopasnost-finansovykh-organizatsiy/ (accessed 20.07.2025) (In Russ.).

14. *How to improve your GOST 57580 score: quickly, effectively, and the first time*. Available at: https://bi.zone/expertise/events/kak-povysit-otsenku-gost-57580-bystro-effektivno-s-pervogo-raza/ (accessed 23.07.2025) (In Russ.).

Статья поступила в редакцию 02.09.2025; одобрена после рецензирования 30.09.2025; принята к публикации 01.10.2025.

The article was submitted 02.09.2025; approved after reviewing 30.09.2025; accepted for publication 01.10.2025.

УДК 004.056

РАЗРАБОТКА ПРОГРАММНОГО РЕШЕНИЯ ДЛЯ АНАЛИЗА ЗАЩИЩЕННОСТИ СРЕДСТВ БЕЗОПАСНОГО УДАЛЕННОГО ПОДКЛЮЧЕНИЯ

Зюбин Владислав Артемович, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
студент, ORCID: 0000-0003-3814-8740, e-mail: vlad.zu2011@mail.ru
Макарян Александр Самвелович, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
кандидат технических наук, доцент, ORCID: 0000-0002-1801-6137, e-mail: makaryan@kubstu.ru
Путято Михаил Михайлович, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
доцент, ORCID: 0000-0003-0414-6034, e-mail: michael.putyato@kubstu.ru
Черкасов Александр Николаевич, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
кандидат технических наук, доцент, ORCID: 0000-0002-5015-4556, e-mail: cherk@mail.ru
Косогорова Маргарита Евгеньевна, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,
бакалавр, ORCID: 0009-0007-0636-3944, e-mail: m-kosogorova@mail.ru

В условиях роста популярности удаленной работы и увеличения числа кибератак на средства удаленного подключения обеспечение их защищенности становится критически важной задачей. В статье рассматриваются ключевые аспекты безопасности удаленного доступа, включая анализ существующих угроз и уязвимостей. Представлено теоретическое обоснование методов защиты, таких как шифрование данных, аутентификация пользователей и контроль доступа. На основе проведенного анализа разработано программное решение, предназначенное для оценки уровня защищенности систем удаленного подключения. Описаны основные функции решения, включая сканирование уязвимостей, анализ конфигураций и мониторинг сетевой активности. Реализация решения включает интеграцию современных методов защиты и инструментов анализа. В рамках тестирования проведена оценка эффективности предложенного решения, подтвердившая его способность выявлять и предотвращать потенциальные угрозы. В результате работы сформулированы рекомендации по повышению уровня защищенности удаленных подключений и предложены пути дальнейшего развития программного решения.

Ключевые слова: кибербезопасность, угроза, уязвимость, удаленное подключение

Финансирование: *исследование выполнено за счет гранта Российского научного фонда № 25-21-00379.*

DEVELOPMENT OF A SOFTWARE SOLUTION FOR ANALYZING THE SECURITY OF REMOTE SECURE CONNECTION FACILITIES

Zyubin Vladislav A., Kuban State Technological University, 2, Moskovskaya St., Krasnodar, 350072, Russian Federation,
student, ORCID: 0000-0003-3814-8740, e-mail: vlad.zu2011@mail.ru
Makaryan Alexander S., Kuban State Technological University, 2, Moskovskaya St., Krasnodar, 350072, Russian Federation,
Cand. Sci. (Engineering), Assistant Professor, ORCID: 0000-0002-1801-6137, e-mail: makaryan @kubstu.ru
Putyato Mikhail M., Kuban State Technological University, 2, Moskovskaya St., Krasnodar, 350072, Russian Federation,
Assistant Professor, ORCID: 0000-0003-0414-6034, e-mail: michael.putyato@kubstu.ru
Cherkasov Alexander N., Kuban State Technological University, 2, Moskovskaya St., Krasnodar, 350072, Russian Federation,
Cand. Sci. (Engineering), Assistant Professor, ORCID: 0000-0002-5015-4556, e-mail: cherk@mail.ru
Kosogorova Margarita E., Kuban State Technological University, 2, Moskovskaya St., Krasnodar, 350072, Russian Federation,
bachelor, ORCID: 0009-0007-0636-3944, e-mail: m-kosogorova@mail.ru

With the growing popularity of remote work and the increasing number of cyber attacks on remote connection facilities, ensuring their security is becoming a critical task. The article discusses key aspects of remote access security, including an analysis of existing threats and vulnerabilities. A theoretical justification of security methods such as data encryption, user authentication, and access control is presented. Based on the analysis, a software solution has been developed to assess the security level of remote connection systems. The main functions of the solution are described, including vulnerability scanning, configuration analysis, and network activity monitoring. The implementation of the

solution includes the integration of modern security methods and analysis tools. As part of the testing, the effectiveness of the proposed solution was evaluated, confirming its ability to identify and prevent potential threats. As a result of the work, recommendations were formulated to increase the security level of remote connections and proposed ways to further develop the software solution.

Keywords: cybersecurity, threat, vulnerability, remote connection

Financial support: *the work was supported by the Russian Science Foundation, Project № 25-21-00379.*

ВВЕДЕНИЕ

В современном мире, где информационные технологии развиваются с невероятной скоростью и все глубже проникают в бизнес-процессы организаций, обеспечивая оптимизацию, модернизацию и повышение эффективности работы, вопросы информационной безопасности становятся ключевыми для успешного функционирования любой компании. Особую актуальность приобретает обеспечение безопасности удаленного подключения, которое является неотъемлемой частью современных корпоративных сетей. С ростом технологического прогресса и расширением возможностей удаленной работы увеличивается и количество угроз, связанных с утечкой данных, несанкционированным доступом и другими киберрисками. В таких условиях разработка программного решения для анализа защищенности средств безопасного подключения становится не просто актуальной, а жизненно необходимой, особенно для организаций, работающих с большими объемами конфиденциальной информации.

Внедрение специализированного программного обеспечения для анализа защищенности удаленных подключений позволит организациям не только оперативно выявлять и устранять уязвимости в системах безопасности, но и проактивно предотвращать потенциальные угрозы. Это особенно важно в условиях, когда киберугрозы становятся все более сложными и изощренными, а их последствия – все более разрушительными. Разработка такого решения будет способствовать не только повышению уровня защищенности корпоративных данных, но и обеспечению соответствия требованиям международных стандартов и нормативных актов, что, в свою очередь, укрепит доверие клиентов и партнеров к организации. Таким образом, создание программного решения для анализа защищенности средств безопасного подключения является важным шагом на пути к построению устойчивой и надежной системы информационной безопасности в условиях цифровой трансформации.

АНАЛИЗ ПРОБЛЕМЫ БЕЗОПАСНОСТИ УДАЛЕННОГО ПОДКЛЮЧЕНИЯ

В условиях стремительного развития информационных технологий и массового перехода на удаленные форматы работы обеспечение безопасности удаленного подключения становится одной из ключевых задач в области информационной безопасности. Удаленный доступ к корпоративным сетям, облачным сервисам и критически важным данным требует надежной защиты от постоянно растущих киберугроз.

Рост числа атак на интернет-соединения, эксплуатация уязвимостей в протоколах удаленного доступа и использование методов социальной инженерии подчеркивают необходимость разработки эффективных решений для анализа и обеспечения защищенности таких систем. По данным исследований, более 60 % утечек данных происходят из-за недостаточной защищенности удаленных подключений, что приводит к значительным финансовым и репутационным потерям.

Особую актуальность теме придает широкое распространение облачных технологий, интернета вещей (IoT) и мобильных устройств, которые расширяют периметр корпоративных сетей и создают новые векторы для атак. Традиционные методы безопасности зачастую оказываются недостаточно эффективными, что требует разработки специализированных программных решений для оперативного выявления уязвимостей и анализа угроз.

Таким образом, проблема обеспечения безопасности удаленного доступа является актуальной как с теоретической, так и с практической точки зрения, что подтверждается необходимостью внедрения подобных решений в различных отраслях, включая финансы, здравоохранение и государственный сектор.

ТЕОРЕТИЧЕСКОЕ ОБОСНОВАНИЕ МЕТОДОВ ЗАЩИТЫ УДАЛЕННОГО ПОДКЛЮЧЕНИЯ

При проектировании системы безопасного удаленного подключения необходимо сформулировать и обеспечить выполнение комплекса требований, направленных на защиту информации и обеспечение надежного функционирования системы. Данные требования должны соответствовать современным стандартам информационной безопасности и учитывать специфику конкретной организации.

Основополагающим требованием является обеспечение конфиденциальности передаваемых данных. Это достигается путем применения современных криптографических протоколов для шифрования трафика. При выборе криптографических алгоритмов следует отдавать предпочтение проверенным временем решениям, имеющим достаточную криптографическую стойкость.

Система должна обеспечивать надежную аутентификацию пользователей при установлении удаленного подключения. Рекомендуется использовать многофакторную аутентификацию, сочетающую различные факторы: пароли, аппаратные токены, биометрические данные. Важным требованием является регулярная смена паролей и применение политик, обеспечивающих их сложность.

Необходимо реализовать механизмы контроля доступа, позволяющие ограничивать права удаленных пользователей в соответствии с их должностными обязанностями. Система должна поддерживать ролевую модель доступа и обеспечивать принцип минимальных привилегий.

Существенным требованием является обеспечение целостности передаваемых данных. Система должна гарантировать, что информация не была изменена при передаче по каналам связи. Для этого следует использовать механизмы электронной подписи и контрольных сумм.

Важным аспектом является обеспечение доступности системы удаленного подключения. Необходимо предусмотреть резервирование каналов связи и ключевых компонентов системы, а также механизмы балансировки нагрузки для обеспечения стабильной работы при большом количестве одновременных подключений.

РАЗРАБОТКА ПРОГРАММНОГО РЕШЕНИЯ ДЛЯ АНАЛИЗА ЗАЩИЩЕННОСТИ

Python выбран в качестве основного языка разработки благодаря простому синтаксису, высокой скорости создания приложений и широкой экосистеме библиотек. Его кроссплатформенность позволяет запускать разработанный инструмент на различных операционных системах, а встроенная поддержка многопоточности через модуль threading обеспечивает эффективное выполнение параллельных задач.

Для реализации функций сетевого сканирования использована библиотека python-nmap, которая предоставляет удобный интерфейс к Nmap – индустриальному стандарту для анализа сетей. Это позволяет с высокой точностью определять версии служб и операционных систем, а также расширять функциональность с помощью существующих NSE-скриптов.

Низкоуровневое взаимодействие с сетевыми протоколами обеспечивается библиотекой scapy. Она позволяет перехватывать, анализировать и создавать кастомизированные сетевые пакеты, что особенно полезно при реализации специализированных функций сканирования.

Для работы с протоколом SMB используется библиотека smbprotocol, обеспечивающая полную поддержку протоколов SMB2/SMB3, что важно для взаимодействия с файловыми и сетевыми ресурсами в корпоративных сетях.

Разработанное программное решение представляет собой модульную систему для комплексного анализа защищенности сетевой инфраструктуры. Архитектура приложения построена с учетом современных принципов разработки программного обеспечения и состоит из следующих основных компонентов.

Модуль управления сканированием отвечает за координацию работы всех остальных модулей. Он принимает входные данные от пользователя (например, диапазон IP-адресов, порты для сканирования, учетные данные для SMB и т. д.) и распределяет задачи между другими модулями.

Модуль ARP-сканирования использует протокол ARP для определения MAC-адресов устройств в локальной сети. Он отправляет ARP-запросы на указанные IP-адреса и анализирует ответы, чтобы определить активные устройства.

Модуль Nmap-сканирования использует библиотеку Nmap для сканирования открытых портов, определения служб и операционных систем на целевых устройствах. Он также выполняет проверку на наличие известных уязвимостей с использованием скриптов Nmap.

Модуль проверки уязвимостей анализирует данные, полученные от модуля Nmap-сканирования, и сравнивает их с базой данных уязвимостей (CVE). Он выявляет потенциальные уязвимости в службах и операционных системах.

Модуль перечисления SMB-ресурсов использует библиотеку SMB для подключения к устройствам с открытым портом 445 и перечисления доступных SMB-ресурсов. Он также проверяет возможность анонимного доступа к ресурсам.

Модуль проверки DNS Zone Transfer проверяет возможность передачи DNS-зоны на устройствах с открытым портом 53. Он отправляет запросы на передачу зоны и анализирует ответы.

Модуль отчетности формирует отчеты о результатах сканирования в текстовом и PDF-формате. Отчеты включают информацию об обнаруженных устройствах, открытых портах, службах, операционных системах и уязвимостях.

Модуль многопоточности позволяет одновременно сканировать несколько IP-адресов, что значительно сокращает время выполнения задачи.

Архитектура сканера обеспечивает высокую гибкость и масштабируемость. Каждый модуль может быть легко заменен или расширен без необходимости изменения других компонентов системы. Схема взаимодействия модулей показана на рисунке 1.

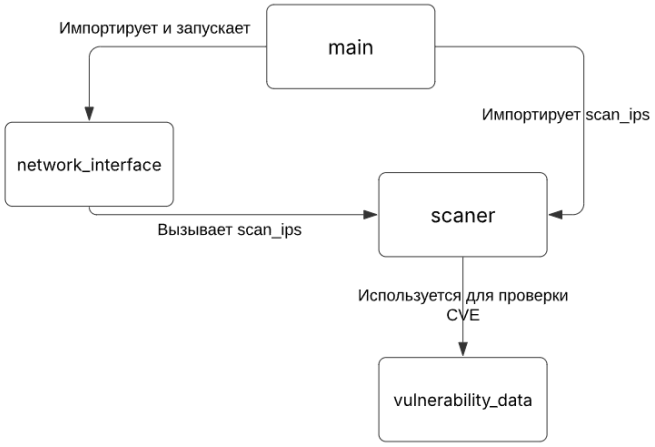


Рисунок 1 – Схема взаимодействия модулей

РЕАЛИЗАЦИЯ ФУНКЦИЙ АНАЛИЗА ЗАЩИЩЕННОСТИ

Реализация функций анализа защищенности является ключевым этапом разработки сетевого сканера, обеспечивающего оценку уровня безопасности сетевых устройств. Основная задача данных функций заключается в идентификации открытых портов, предоставляемых сервисов, а также в выявлении известных уязвимостей на основе данных о CVE (Common Vulnerabilities and Exposures). Разработанный инструмент включает несколько взаимосвязанных функций, которые совместно обеспечивают комплексный анализ защищенности.

Функция perform_nmap_scan использует библиотеку Nmap для сканирования заданных портов и сбора информации об открытых портах, сервисах и операционной системе. Она инициализирует сканер Nmap, выполняет сканирование указанного IP-адреса и портов, а затем извлекает и структурирует данные о состоянии портов, версиях сервисов и операционной системе. Это позволяет получить детальную информацию о сетевых устройствах, что является основой для дальнейшего анализа защищенности.

Функция check_vulnerabilities проверяет информацию о сервисах и операционной системе на наличие известных уязвимостей, используя базу данных CVE. Для этого данные CVE загружаются из JSON-файла, после чего строится структура для быстрого поиска уязвимостей. Каждый сервис и операционная система проверяются на соответствие с записями в базе данных CVE, что позволяет выявить потенциальные уязвимости. Это особенно важно для оценки рисков, связанных с использованием устаревших или небезопасных версий программного обеспечения.

Функция enumerate_smb_shares использует библиотеку SMBConnection для перечисления SMB-ресурсов на заданном IP-адресе. Это позволяет оценить безопасность SMB-сервисов, которые часто являются целью атак из-за их широкого распространения и потенциальных уязвимостей. Перечисление ресурсов помогает выявить открытые для доступа файловые системы, что может представлять серьезную угрозу для безопасности сети.

Функция check_dns_zone_transfer проверяет возможность передачи DNS-зоны с использованием библиотеки dns.query. DNS-зона содержит информацию о доменных именах и соответствующих IP-адресах, и ее несанкционированная передача может привести к утечке конфиденциальных данных. Проверка этой функции позволяет выявить неправильно настроенные DNS-серверы, которые могут быть использованы злоумышленниками.

Функция scan_single_ip координирует выполнение вышеперечисленных функций для сканирования одного IP-адреса, собирая всю необходимую информацию для анализа защищенности. Она объединяет результаты ARP-сканирования, Nmap-сканирования, проверки уязвимостей и других операций, предоставляя полную картину состояния безопасности устройства. Эти функции интегрированы в основной процесс сканирования, который выполняется в функции scan_ips. Данная функция разбирает диапазон IP-адресов, использует многопоточность для одновременного сканирования нескольких IP-адресов и собирает результаты сканирования для каждого из них. Это позволяет эффективно анализировать большие сети за минимальное время.

ТЕСТИРОВАНИЕ И ОЦЕНКА ЭФФЕКТИВНОСТИ РЕШЕНИЯ

Тестирование разработанного программного обеспечения представляет собой ключевой этап, который позволяет оценить его работоспособность, корректность выполнения функций и соответствие установленным требованиям. Главная цель тестирования заключается в проверке функциональности системы, выявлении потенциальных ошибок и уязвимостей, а также в оценке эффективности работы системы в различных условиях.

Функциональное тестирование – проверка корректности работы всех функций системы, таких как сканирование IP-адресов, определение открытых портов, проверка уязвимостей и формирование отчетов. В ходе функционального тестирования были проверены все основные функции системы. Система корректно определяет активные устройства в сети с использованием ARP-сканирования, успешно выявляет открытые порты и службы на устройствах с помощью Nmap-сканирования, а также корректно идентифицирует уязвимости на основе данных из базы CVE. Кроме того, система формирует отчеты в формате PDF, которые содержат всю необходимую информацию, включая IP-адреса, MAC-адреса, открытые порты, обнаруженные уязвимости и перечень SMB-ресурсов.

Для проведения тестирования были созданы тестовые сценарии, охватывающие различные аспекты работы системы, включая сканирование одиночных IP-адресов, диапазонов адресов, а также проверку корректности работы в условиях ограниченной сети. Тестирование проводилось на различных операционных системах и сетевых конфигурациях, что позволило убедиться в кроссплатформенной совместимости системы. Результаты тестирования подтвердили, что система стабильно работает в различных условиях и способна эффективно выполнять поставленные задачи.

Тестирование производительности включает оценку времени, затрачиваемого на сканирование сети в различных условиях (количество IP-адресов, количество портов, наличие SMB-ресурсов). Система успешно справляется с задачей сканирования сетей, охватывая диапазон от 10 до 255 IP-адресов. Время сканирования увеличивается пропорционально росту числа адресов. Кроме того, система эффективно сканирует до 1000 портов на каждом устройстве, и, несмотря на увеличение времени сканирования с ростом числа портов, оно остается приемлемым для практического использования. Также система корректно определяет и проверяет доступность SMB-ресурсов.

В таблице и на рисунке 2 наглядно представлена производительность.

Таблица – Производительность

Количество IP-адресов	Количество портов	Время сканирования (без многопоточности), с	Время сканирования (с многопоточностью), с
10	50	120	30
20	100	240	60
50	200	600	150
100	500	1200	300

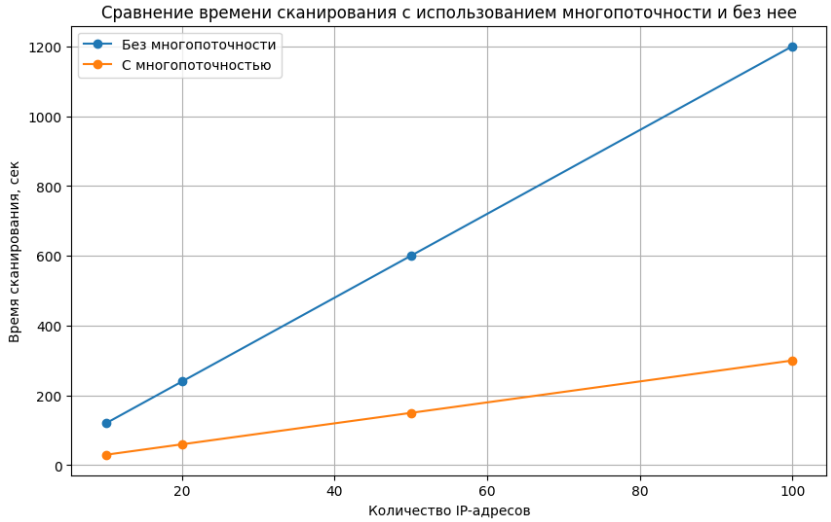


Рисунок 2 – График производительности

Тестирование устойчивости – проверка работы системы при наличии сетевых ошибок, таких как недоступность некоторых узлов сети или высокая загрузка сети.

Тестирование безопасности – проверка корректности работы системы при использовании SMB-аутентификации и других механизмов безопасности. Система корректно обрабатывает недоступные узлы сети, продолжая сканирование остальных устройств без прерывания. Кроме того, она сохраняет работоспособность даже в условиях высокой загрузки сети, хотя в таких случаях время сканирования может увеличиваться.

В результате тестирования было подтверждено, что разработанная система эффективно выполняет свои функции, обладает высокой производительностью и устойчивостью к различным сетевым условиям. Система может быть использована для автоматизации процессов сканирования сети, определения уязвимостей и формирования отчетов, что делает ее полезным инструментом для специалистов в области информационной безопасности.

Список источников

1. Зюбин, В. А. Дипломная работа по специальности 10.05.01 «Компьютерная безопасность» / В. А. Зюбин. – Краснодар : Кубанский государственный технический университет, 2025. – 66 с.
2. Федеральный закон Российской Федерации от 27 июля 2006 г. № 152-ФЗ «О персональных данных» (ред. 31 декабря 2017 г.).
3. Федеральный закон Российской Федерации от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации» (ред. 1 марта 2019 г.).
4. ГОСТ Р 50922-2006 «Защита информации. Основные термины и определения».
5. Постановление Правительства Российской Федерации от 15.09.2008 № 687 «Об утверждении Положения об особенностях обработки персональных данных, осуществляемой без использования средств автоматизации».
6. Постановление Правительства Российской Федерации от 01.11.2012 № 1119 «Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных». Совместный приказ ФСТЭК России, ФСБ России и Мининформсвязи России от 13.02.2008 № 55/86/20 «Об утверждении порядка проведения классификации информационных систем персональных данных».
7. Приказ ФСТЭК России от 11.02.2013 № 17 «Об утверждении требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах».
8. Бирюков, А. А. Информационная безопасность: защита и нападение / А. А. Бирюков. – Москва : ДМК Пресс, 2017.

References

1. Zyubin, V. A. *Thesis on specialty 10.05.01 "Computer security".* Krasnodar, Kuban State Technological University, 2025. 66 p. (In Russ.).
2. *Federal Law of the Russian Federation of July 27, 2006 no. 152-FZ "On Personal Data" (as amended on December 31, 2017)* (In Russ.).
3. *Federal Law of the Russian Federation of July 27, 2006 no. 149-FZ "On Information, Information Technologies and Information Protection" (as amended on March 1, 2019)* (In Russ.).
4. *GOST R 50922-2006 "Information protection. Basic terms and definitions"* (In Russ.).
5. *Decree of the Government of the Russian Federation dated 09/15/2008 no. 687 "On Approval of the Regulation on the Specifics of personal Data processing carried out without the use of automation tools"* (In Russ.).
6. *Decree of the Government of the Russian Federation dated 01.11.2012 no. 1119 "On Approval of Requirements for the Protection of Personal Data during their Processing in Personal Data Information Systems". Joint Order of the FSTEC of Russia, the FSB of Russia and the Ministry of Information and Communications of Russia dated 13.02.2008 no. 55/86/20 "On Approval of the Procedure for Classifying Personal Data Information Systems"* (In Russ.).
7. *Order of the FSTEC of Russia dated 11.02.2013 no. 17 "On Approval of Requirements for the Protection of Information not Constituting a State Secret contained in State Information Systems"* (In Russ.).
8. Biryukov, A. A. *Information security: protection and attack*. Moscow, DMK Press, 2017 (In Russ.).

Статья поступила в редакцию 11.09.2025; одобрена после рецензирования 30.09.2025; принята к публикации 14.11.2025.

The article was submitted 11.09.2025; approved after reviewing 30.09.2025; accepted for publication 14.11.2025.

УДК 004.6

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ СКРЫТОЙ МАРКИРОВКИ ДОКУМЕНТОВ НА ОСНОВЕ СТЕГАНОГРАФИИ

Вишневский Андрей Сергеевич, Астраханский государственный университет им. В. Н. Татищева, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
аспирант, e-mail: vishnevskiy1.16@gmail.com

В статье рассматриваются современные методы скрытой маркировки документов с использованием стеганографии, направленные на защиту конфиденциальной информации и контроль за распространением копий. Проведен сравнительный анализ классических алгоритмов (Bruyndonckx, Pitas, Koch, Cox, Barni, Wang) и современных подходов, основанных на глубоких автоэнкодерах, обучении с подкреплением и гибридных методах LSB-DWT с энтропийным шифрованием. Основные критерии оценки включают незаметность внедрения, устойчивость к сжатию, фильтрации и геометрическим преобразованиям, емкость и вычислительную сложность. Результаты показывают, что классические методы обеспечивают высокую точность восстановления в ограниченных условиях, однако имеют низкую универсальность, тогда как современные алгоритмы демонстрируют сбалансированную устойчивость ко всем видам атак и сохраняют визуальную идентичность документов. Отдельное внимание уделяется текстовым документам, где эффективны методы внедрения невидимых символов и корректировки форматирования, а для изображений и PDF-документов – частотные и комбинированные подходы. В работе также подчеркивается важность интеграции организационных мер, включая аудит и логирование, для комплексной защиты данных. Результаты исследования будут в дальнейшем использованы при разработке интеллектуальных систем защиты корпоративной информации и цифровой сертификации документов.

Ключевые слова: стеганография, скрытая маркировка, LSB, DCT, DWT, устойчивость, цифровая водяная метка, цифровой водяной знак

COMPARATIVE ANALYSIS OF METHODS OF HIDDEN MARKING OF DIGITAL DOCUMENTS BASED ON STEGANOGRAPHY

Vishnevsky Andrey S., Astrakhan Tatishchev State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
postgraduate student, e-mail: vishnevskiy1.16@gmail.com

The article discusses modern methods of covert labeling of documents using steganography, aimed at protecting confidential information and controlling the distribution of copies. A comparative analysis of classical algorithms (Bruyndonckx, Pitas, Koch, Cox, Barni, Wang) and modern approaches based on deep autoencoders, reinforcement learning, and hybrid LSB-DWT methods with entropy encryption is carried out. The main evaluation criteria include invisibility of embedding, resistance to compression, filtration and geometric transformations, capacity and computational complexity. The results show that classical methods provide high accuracy of recovery in limited conditions, but have low versatility, while modern algorithms demonstrate balanced resistance to all types of attacks and preserve the visual identity of documents. Special attention is paid to text documents, where methods of introducing invisible characters and correcting formatting are effective, and frequency and combined approaches are used for images and PDF documents. The paper also highlights the importance of integrating organizational measures, including auditing and logging, for comprehensive data protection. The research results can be used in the development of intellectual systems for the protection of corporate information and digital certification of documents.

Keywords: steganography, hidden marking, LSB, DCT, DWT, robustness, digital watermark

ВВЕДЕНИЕ

В условиях стремительного роста объемов цифровой информации и активного распространения технологий удаленного доступа проблема защиты конфиденциальных данных приобретает особую актуальность. Современные организации сталкиваются с угрозами несанкционированного копирования, модификации и утечки служебных документов, что может привести к существенным финансовым потерям и подрыву репутации компаний. В 2024 г. было зафиксировано более 1,7 миллиарда случаев компрометации личных данных, что является значительным увеличением по сравнению с предыдущим годом [1]. Из них шесть крупных инцидентов затронули более 100 миллионов записей каждый, включая утечку данных от Ticketmaster, затронувшую 560 миллионов человек [2].

Одним из эффективных направлений в обеспечении информационной безопасности является использование методов скрытой маркировки, позволяющих внедрять в документ невидимые идентификационные атрибуты, по которым впоследствии можно определить источник утечки или подтвердить подлинность копии.

Скрытая маркировка рассматривается как частный случай стеганографических методов, ориентированных на сокрытие информации внутри других цифровых данных таким образом, чтобы факт присутствия скрытого сообщения оставался незаметным. То есть в отличие от классических

криптографических подходов, стеганография не только защищает содержание данных, но и скрывает сам факт их передачи или хранения. Это делает ее особенно ценной в системах, где требуется контроль распространения документов при сохранении их внешнего вида и структуры без видимых следов вмешательства.

Среди множества направлений стеганографии особое место занимают методы, применяемые для маркировки текстовых, графических и комбинированных документов. Для каждого типа данных используются различные принципы внедрения – от изменения форматирования и внедрения невидимых символов в текст до модификации наименее значимых бит изображения или частотных коэффициентов при сжатии. Такие методы позволяют создавать устойчивые, незаметные и уникальные метки, обеспечивающие контроль за распространением копий и возможность последующей верификации.

Целью данной работы является проведение сравнительного анализа существующих алгоритмов стеганографии в контексте их использования для скрытой маркировки документов. Рассматриваются подходы к внедрению и сохранению маркировки в процессе копирования, анализируются критерии оценки устойчивости и незаметности стеганографических методов, а также формулируются рекомендации по выбору оптимальных алгоритмов для различных типов цифровых данных.

Переходя к анализу существующих методов, важно рассмотреть основные классы стеганографических алгоритмов, определить их принципы работы и особенности применения в контексте защиты документов, что позволит сформировать теоретическую основу для последующего их сопоставления.

ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ СКРЫТОЙ МАРКИРОВКИ

Современные методы скрытой маркировки цифровых документов основаны на принципах стеганографии – науки о скрытой передаче информации, при которой само существование сообщения остается незаметным для стороннего наблюдателя [3]. Как было ранее сказано, основная цель этих методов – внедрить дополнительную информацию в структуру документа таким образом, чтобы визуальные и структурные свойства исходного объекта оставались практически неизменными. В отличие от видимых водяных знаков, стеганографические метки практически не различимы и могут быть извлечены только специализированным программным обеспечением. Ключевое свойство таких методов – способность сохраняться при копировании, преобразовании формата и частичных изменениях документа, что делает их особенно ценными для защиты конфиденциальной информации.

Методы стеганографии условно делятся на две основные группы: работающие в пространственной области и основанные на преобразованиях в частотной области.

Методы пространственной области

Наиболее известным подходом является метод наименее значимых битов (Least Significant Bit (LSB)) [4]. Суть метода заключается в замене младших битов исходного носителя на биты скрываемого сообщения. Для 8-битного изображения операция внедрения описана формулой:

$$I'(x,y) = I(x,y) - (I(x,y) \bmod 2) + b, \quad (1)$$

где $I(x,y)$ – значение пикселя исходного изображения;

b – бит скрываемого сообщения;

$I'(x,y)$ – значение пикселя после внедрения.

Для цветного изображения LSB необходимо применять по каждому из каналов RGB, что увеличивает емкость хранения.

Для текстовых документов используются аналогичные методы с невидимыми символами (zero-width characters, U+200B, U+200C, U+200D), пробелами с измененной шириной или модификацией кодировки. Например, внедрение последовательности бит $m_i \in \{0,1\}$ в текст может быть реализовано через манипуляцию пробелами:

$$S_i = \begin{cases} \text{обычный пробел, если } m_i = 0 \\ \text{неразрывный пробел, если } m_i = 1 \end{cases}$$

Такая методика позволяет внедрять идентификаторы без изменения визуального восприятия текста, сохраняя читаемость документа.

Методы частотной области

Для повышения устойчивости к преобразованиям применяются методы внедрения в частотную область, основанные на дискретном косинусном преобразовании (DCT) [5] и дискретном вейвлет-преобразовании (DWT) [6]. Для блока 8×8 пикселей DCT внедрение метки описывается формулой:

$$C'(гбм) = C(гбм) + \alpha \cdot ь(гбм)б, \quad (2)$$

где $C(u,v)$ – исходный коэффициент DCT;

$m(u,v)$ – бит сообщения;

α – коэффициент, регулирующий баланс между незаметностью и устойчивостью.

Метка обычно внедряется в коэффициенты средних частот, чтобы минимизировать визуальные искажения.

В DWT сигнал документа разлагается на несколько уровней:

$$I \xrightarrow{DWT} \{A_n, H_n, V_n, D_n\}, \quad (3)$$

где A_n – приближающая компонента;

H_n, V_n, D_n – горизонтальные, вертикальные и диагональные детализированные коэффициенты на уровне n.

Биты сообщения внедряются в выбранные коэффициенты с использованием малых коэффициентов усиления для минимизации визуальных изменений, а распределение метки по нескольким уровням повышает устойчивость к сжатию и шуму.

Комбинированные и адаптивные методы

Комбинированные алгоритмы, такие как гибрид DCT-DWT или Spread Spectrum [7], объединяют преимущества различных подходов. В Spread Spectrum скрытый сигнал *m(t)* распределяется по множеству частот исходного носителя *x(t)*:

$$y(t)=x(t)+\alpha \cdot m(t) \cdot c(t), \quad (4)$$

где *c(t)* – псевдослучайная последовательность для кодирования сигнала. Такой подход повышает устойчивость метки к удалению и модификации.

Современные исследования активно используют глубокие нейронные сети для стеганографии. Алгоритмы обучаются автоматически внедрять данные с минимальными визуальными и статистическими изменениями, сохраняя визуальную и статистическую структуру исходного документа [8–9].

Таким образом, современная стеганография предлагает широкий спектр методов, различающихся по емкости, устойчивости и незаметности внедрения. Выбор алгоритма определяется типом документа и требованиями к защите, что обеспечивает оптимальный баланс между скрытностью и устойчивостью метки.

КРИТЕРИИ ОЦЕНКИ АЛГОРИТМОВ

Эффективность методов скрытой маркировки оценивается по четырем основным критериям: **незаметность**, **устойчивость**, **емкость** и **вычислительная сложность**. Каждый из критериев характеризует разные аспекты работы алгоритма и определяет его практическую применимость для конкретного типа документа.

Незаметность *(imperceptibility)* отражает степень того, насколько внедрение метки влияет на визуальные, структурные и семантические свойства документа. Ключевое требование – сохранение исходного восприятия документа для пользователя и отсутствие заметных изменений.

Для графических и мультимедийных документов оценка выполняется с использованием объективных метрик качества изображения:

1. **PSNR (Peak Signal-to-Noise Ratio)** – показатель отношения сигнала к шуму, измеряемый в децибелах. Для изображения I (оригинал) и I' (с внедренной меткой) PSNR определяется как:

$$PSNR = 10 * log_{10} * \frac{MAX_I^2}{MSE}, \quad (5)$$

где MAX_I – максимальное значение пикселя (для 8-битного изображения $MAX_I = 255$);

MSE – среднеквадратичная ошибка.

2. **SSIM (Structural Similarity Index)** – индекс структурного сходства, учитывающий яркость, контраст и структуру изображения:

$$SSIM(I, I') = \frac{(2\mu_I\mu_{I'}+C_1)(2\sigma II'+C_2)}{(\mu_I{}^2+\mu_{I'}{}^2+C_1)(\sigma_I^2+\sigma_{I'}^2+C_2)}, \quad (6)$$

где μ – среднее значение яркости;

σ^2 – дисперсия;

$\sigma II'$ – ковариация;

C_1 и C_2 – небольшие константы для стабилизации.

Чем выше PSNR и SSIM, тем менее заметна внедренная метка.

Для текстовых документов: незаметность определяется сохранением читаемости, структуры абзацев, форматирования и отсутствием видимых артефактов. Методы с невидимыми символами (пробел нулевой ширины (zero-width characters), модифицированные пробелы, тонкие изменения кодировки) позволяют внедрять данные без визуального изменения текста.

Устойчивость *(robustness)* характеризует способность метки сохраняться при воздействии преобразований и операций, которые могут изменять содержимое документа.

Основные виды воздействия включают:

- копирование и вставку;
- печать и сканирование;
- сжатие (JPEG, PDF, ZIP);
- масштабирование, вращение, обрезку;
- фильтрацию и шум.

Методы частотной области (DCT, DWT) и комбинированные алгоритмы обладают высокой устойчивостью, так как внедряются в средние и низкие частоты, которые меньше подвержены изменениям при сжатии или редактировании.

Метрикой устойчивости является коэффициент корреляции (Correlation Coefficient, *CC*) между исходным и извлеченным сообщением:

$$CC = \frac{\sum_{i=1}^{N}(m_i-\bar{m})(m'_i-\overline{m'})}{\sqrt{\sum_{i=1}^{N}(m_i-\bar{m})^2 \sum_{i=1}^{N}(m'_i-\overline{m'})^2}}, \quad (7)$$

где m_i и m'_i – оригинальные и извлеченные биты сообщения соответственно.

Значение *CC*, близкое к 1, говорит о высокой устойчивости.

***Емкость** (capacity)* отражает объем информации, который можно внедрить в документ без заметных изменений.

Для текстов: количество бит, которое можно скрыть через невидимые символы, пробелы, тонкие изменения кодировки. Обычно ограничено несколькими символами на 1 кБ текста.

Для изображений и мультимедиа: количество бит на пиксель (*bp*), которое можно безопасно использовать. В LSB-методе максимальная емкость = число пикселей × число используемых бит на канал.

Увеличение емкости часто снижает незаметность и устойчивость, поэтому выбор объема внедряемой информации – компромисс между скрытностью и информативностью.

Вычислительная сложность характеризует ресурсоемкость алгоритма: время и память, необходимые для внедрения метки, извлечения, верификации.

Простые LSB-методы обладают низкой сложностью и могут применяться к большим объемам данных, но устойчивость невысока. Методы DCT/DWT и гибридные подходы требуют вычислительных ресурсов для преобразований и обратного восстановления, но обеспечивают большую стойкость метки.

В корпоративных системах с большим объемом документов важно выбирать алгоритмы с оптимальным балансом вычислительной нагрузки и устойчивости, особенно при автоматизации процессов контроля.

Подводя итог, можно сказать, что сопоставление этих критериев позволяет объективно оценивать алгоритмы стеганографии и выбирать оптимальный метод для конкретного типа документа.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Сравнительный анализ алгоритмов стеганографии позволяет выявить преимущества и ограничения каждого метода в контексте скрытой маркировки конфиденциальных документов. Методы, основанные на изменении наименее значимых битов изображений (LSB), обладают высокой емкостью и простотой реализации, однако их устойчивость к редактированию, сжатию и многократному копированию ограничена. В текстовых документах аналогичная зависимость проявляется при использовании невидимых символов и изменений форматирования: они сохраняют визуальную целостность текста, но могут быть легко удалены при конвертации или исправлении форматирования.

Алгоритмы, использующие преобразования в частотной области, такие как DCT и DWT, демонстрируют высокую устойчивость к сжатию, масштабированию и другим преобразованиям, что делает их предпочтительными для графических материалов и PDF-документов с интегрированными изображениями. Комбинированные методы, включая гибридные DWT-DCT и адаптивные алгоритмы с распределением идентификатора по нескольким слоям, обеспечивают оптимальный баланс между стойкостью и незаметностью, минимизируя риск разрушения метки при изменении документа.

Среди наиболее известных классических решений можно выделить алгоритмы **Bruyndonckx, Pitas, Koch, Cox, Barni и Wang**, отличающиеся по области внедрения и типу стегосистемы. Пространственные методы **Bruyndonckx** и **Pitas** применяют прямые изменения яркостных значений пикселей. В первом используется двойной порог ($t_1 = 5$, $t_2 = 10$), что обеспечивает простоту реализации, но низкую устойчивость к компрессии и фильтрации. Алгоритм Pitas выполняет блочное внедрение с размером блока 16×16 пикселей, что повышает устойчивость к локальным искажениям при некотором снижении скрытности.

Метод **Koch** реализует частотный подход на основе дискретного косинусного преобразования (ДКП) с модификацией коэффициентов в блоках 8×8 при уровне квантования 3, демонстрируя хорошее соотношение между визуальной прозрачностью и устойчивостью к JPEG-сжатию.

Класс закрытых стегосистем представлен алгоритмами **Cox, Barni и Wang**, использующими числовые последовательности в качестве цифрового водяного знака. Алгоритм Cox внедряет псевдослучайную последовательность в коэффициенты ДКП размером n×n ($n = 100$), обеспечивая высокую робастность к сжатию и масштабированию. Метод Barni переносит внедрение в область вейвлет-преобразования, где изменение коэффициентов подстраивается под локальные характеристики изображения, обеспечивая адаптивную устойчивость. Алгоритм Wang развивает подход Barni, вводя

параметр $\beta = 1$, позволяющий регулировать амплитуду внедрения в зависимости от энергии поддиапазона, что повышает устойчивость при сохранении высокого значения PSNR (более 45 дБ).

В последние годы наблюдается переход от классических методов к подходам, основанным на **глубоком обучении и нейросетевых моделях**. Одним из таких решений является **RoSteALS (Robust Steganography using Autoencoder Latent Space, 2023)**, в котором внедрение данных осуществляется в латентное пространство автоэнкодера. Такой подход обеспечивает высокую устойчивость к сжатию и фильтрации, а также «слепое» извлечение информации без исходного контейнера [10].

Метод **Deep Convolutional Autoencoder Steganography** на основе **ResNet-архитектуры (2022)** позволяет внедрять одно изображение в другое с сохранением визуальной идентичности контейнера (PSNR > 40 дБ, SSIM > 0,98) [11]. Эти решения используют преимущества глубоких сверточных сетей, способных обучаться оптимальному распределению модификаций пикселей для максимальной скрытности.

Дополнительное направление представлено **Deep Reinforcement Learning-Based DCT Steganography (2023)**, где выбор областей внедрения осуществляется с помощью обучения с подкреплением. Это позволяет адаптировать процесс стеганографирования под конкретные характеристики изображения, снижая вероятность обнаружения и повышая устойчивость.

В последние годы также активно исследуются **многоуровневые гибридные подходы**, объединяющие классические методы (например, LSB или DWT) с механизмами глубокого кодирования и энтропийного шифрования. Такие решения обеспечивают более высокий уровень безопасности за счет распределенного внедрения и динамического выбора областей скрытия.

Таким образом, современное развитие стеганографических технологий направлено на интеграцию классических алгоритмов с нейросетевыми архитектурами, что позволяет достигать баланса между скрытностью, устойчивостью и вычислительной эффективностью. Подобные методы демонстрируют высокую применимость для задач защиты интеллектуальной собственности, цифровой сертификации и маркировки мультимедийных документов в условиях динамических преобразований данных.

Для оценки эффективности алгоритмов стеганографии была проведена сравнительная характеристика как классических, так и современных методов скрытой маркировки документов. Основными критериями оценки выступали однозначность восстановления скрытого сообщения, устойчивость к фильтрации, геометрическим преобразованиям и сжатию, а также стойкость к статистическому стегоанализу.

В таблице представлены результаты исследования стойкости различных алгоритмов, включающих как традиционные подходы (Bruyndonckx, Pitas, Koch, Cox, Barni, Wang), так и современные методы, основанные на глубоких нейронных сетях, обучении с подкреплением и комбинированных подходах LSB-DWT. Классические алгоритмы демонстрируют определенную эффективность в конкретных сценариях: пространственные методы хорошо сохраняют однозначность восстановления, но чувствительны к редактированию и сжатию, а алгоритмы частотной области обеспечивают устойчивость к обработке изображения, но ограничены по емкости и вычислительной сложности.

Современные методы, такие как RoSteALS, сверточные автоэнкодеры, Deep RL-DCT и комбинированные подходы LSB+DWT с энтропийным шифрованием, показывают высокий уровень устойчивости ко всем видам атак, одновременно обеспечивая точное восстановление ЦВЗ и сохранение визуальной целостности документа. Эти алгоритмы демонстрируют способность автоматически адаптироваться к характеристикам контейнера и корректно извлекать скрытую информацию даже после сжатия, фильтрации или геометрических преобразований.

Однозначность восстановления отражает способность извлечь ЦВЗ без ошибок. Устойчивость к фильтрации, сжатию и геометрическим преобразованиям оценивается на основе воздействия стандартных атак (фильтры низких частот, JPEG-сжатие, повороты и масштабирование). Статистический стегоанализ оценивает вероятность обнаружения скрытой информации при анализе распределения битовых значений.

Представленные данные позволили выявить сильные и слабые стороны каждого метода и служат основой для выбора оптимального алгоритма в зависимости от типа документа, требуемого уровня защиты и практических требований к внедрению маркировки.

Анализ показал, что для текстовых документов наиболее эффективными являются методы невидимых символов с уникальными идентификаторами и контролем форматирования, тогда как для изображений и PDF-документов целесообразно использовать частотные и комбинированные подходы. В дополнение к алгоритмическим характеристикам практическая применимость определяется интеграцией с системами верификации, автоматическим восстановлением маркировки при копировании и наличием многоуровневой защиты от удаления.

Таблица – Сравнительные характеристики при исследовании стойкости алгоритмов стеганографии

Название алгоритма	**Тип стегосистемы**	**ЦВЗ**	**Область преобразования**	**Однозначность восстановления**	**Устойчивость к фильтрации**	**Устойчивость к геометрическим преобразованиям**	**Устойчивость к сжатию**	**Устойчивость к статистическому стегоанализу**
Bruyndonckx	Открытая	Текстовая строка	Пространственная	+	–	–	–	+
Pitas	Открытая	Текстовая строка	Пространственная	+	–	–	+	+
Koch	Открытая	Текстовая строка	Блочное ДКП 8×8	+	–	–	–	–
Cox	Закрытая	Числовая последовательность	Блочное ДКП n×n, n = 100	+	–	+	+	–
Barni	Закрытая	Числовая последовательность	Вейвлет-преобразование, n = 100	–	+	–	+	+
Wang	Закрытая	Числовая последовательность	Вейвлет-преобразование, n = 100, β = 1	+	+	–	+	–
RoSteALS	Закрытая	Бинарный код	Латентное пространство автоэнкодера	+	+	+	+	+
Deep Convolutional Autoencoder	Закрытая	Бинарный код	Скрытие в изображении с использованием сверточных автоэнкодеров	+	+	+	+	+
Deep RL-Based DCT	Закрытая	Числовая последовательность	DCT с адаптивным выбором областей через обучение с подкреплением	+	+	+	+	+
LSB+DWT (энгтропийное шифрование)	Закрытая	Числовая последовательность	Комбинированная LSB и вейвлет-преобразование	+	+	+	+	+

Примечание. «+» – высокая устойчивость /точность восстановления; «–» – низкая устойчивость / точность восстановления.

Сравнительный анализ показал, что универсального метода для всех типов документов на данный момент не выявлено, и эффективная система маркировки должна комбинировать различные подходы с учетом специфики данных и организационных требований.

ЗАКЛЮЧЕНИЕ

Скрытая маркировка конфиденциальных документов представляет собой эффективный инструмент контроля за распространением информации и обеспечения ее целостности. Проведенный анализ методов стеганографии показал, что выбор алгоритма должен основываться на типе данных, требуемой степени устойчивости и незаметности внедрения, а также на возможности автоматизации процессов верификации и восстановления маркировки. Для текстовых документов предпочтительны методы внедрения невидимых символов и корректировки форматирования, обеспечивающие сохранение читаемости и уникальность идентификаторов. Для графических материалов и PDF-документов более эффективны алгоритмы частотной обработки и комбинированные подходы, позволяющие достигать высокой устойчивости к редактированию, сжатию и многократному копированию.

Сравнительный анализ классических алгоритмов стеганографии (Bruyndonckx, Pitas, Koch, Cox, Barni, Wang) показал, что ни один из них не обеспечивает универсальной защиты во всех условиях. Алгоритмы Bruyndonckx и Pitas демонстрируют высокую точность восстановления при минимальных искажениях, однако чувствительны к геометрическим трансформациям. Метод Cox, использующий числовую последовательность и блочное ДКП-преобразование, показал высокую устойчивость к сжатию, но низкую – к фильтрации. Наиболее сбалансированные результаты демонстрирует метод Barni, основанный на вейвлет-преобразовании и распределении цифрового водяного знака по нескольким поддиапазонам частот, что обеспечивает устойчивость к фильтрации и статистическому анализу.

Современные исследования развивают данные подходы в направлении **адаптивных и интеллектуальных стеганографических систем.** В частности, **алгоритмы на основе сингулярного разложения матриц (SVD)** обеспечивают устойчивость к сжатию JPEG и масштабированию за счет внедрения ЦВЗ в сингулярные значения, устойчивые к линейным преобразованиям. **Методы на основе дискретного косинусного преобразования с нейронной оптимизацией (DeepDCT)** используют сверточные нейронные сети для динамического выбора областей внедрения, что повышает незаметность и стойкость к статистическому анализу.

Дополнительно **модели на основе глубоких автоэнкодеров (Deep Steganography)**, реализующие скрытие данных на уровне пиксельных распределений, позволяют кодировать большие объемы информации с сохранением визуальной идентичности контейнера. Эти методы демонстрируют высокий уровень автоматической адаптации к типу носителя и устойчивость к перекодированию, что делает их перспективными для применения в интеллектуальных системах защиты корпоративных данных.

Кроме того, важно отметить, что комплексная защита информации невозможна без организационных мер, включающих логирование операций, регулярный аудит и обучение сотрудников, что позволит поддерживать высокий уровень информационной безопасности. Многослойные и криптографически защищенные методы маркировки, интегрированные с системой контроля и аудита, создадут надежный барьер для несанкционированного доступа и позволят отслеживать источники утечек.

Таким образом, эффективная система скрытой маркировки документов строится на сочетании **классических, современных алгоритмических и организационных мер**, обеспечивая уникальность и стойкость внедренных идентификаторов, контроль за копированием и распространением информации, а также возможность быстрой верификации подлинности документов. Такой подход обеспечивает комплексную защиту информации, снижает риски утечки и формирует основу для дальнейшего развития методов стеганографии и интеграции их в корпоративные системы безопасности.

Список источников

1. HIPAA Journal. More Than 1.7 Billion Individuals Had Personal Data Compromised in 2024. – URL: https://www.hipaajournal.com/1-7-billion-individuals-data-compromised-2024/ (дата обращения: 18.09.2025).

2. Watts, L. The 7 Most Telling Data Breaches of 2024. Nightfall / L. Watts. – 2024. – URL: https://www.nightfall.ai/blog/the-7-most-telling-data-breaches-of-2024 (дата обращения: 18.09.2025).

3. Стеганография // Security Vision. – URL: https://www.securityvision.ru/education/cyberwiki/s-rus/steganografiya/ (дата обращения: 20.09.2025).

4. Least Significant Bit (LSB) // Tuple Knowledge Base. – URL: https://www.tuple.nl/en/knowledge-base/least-significant-bit-lsb (дата обращения: 22.09.2025).

5. Discrete Cosine Transform (DCT) // Exponenta Docs. – URL: https://docs.exponenta.ru/images/discrete-cosine-transform.html (дата обращения: 22.09.2025).

6. Wavelet Transform // Gwyddion User Guide. – URL: https://gwyddion.net/documentation/user-guide-ru/wavelet-transform.html (дата обращения: 28.09.2025).

7. CNSS Instructions // Committee on National Security Systems. – URL: https://www.cnss.gov/CNSS/issuances/Instructions.cfm (дата обращения: 28.09.2025).

8. Jiren, Zhu. HiDDeN: Hiding Data With Deep Networks / Jiren Zhu, Russell Kaplan, Justin Johnson, and Li Fei-Fei // Computer Vision – ECCV 2018 : 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part XV. – Berlin, Heidelberg : Springer-Verlag, 2018. – P. 682–697. https://doi.org/10.1007/978-3-030-01267-0_40.

9. Himthani, V. Comparative performance assessment of deep learning based image steganography techniques / V. Himthani, V. S. Dhaka, M. Kaur et al. // Sci. Rep. – 2022. – Vol. 12. – P. 16895. https://doi.org/10.1038/s41598-022-17362-1.

10. Bui, T. RoSteALS: Robust Steganography using Autoencoder Latent Space / T. Bui, S. Agarwal, N. Yu, & J. Collomosse // arXiv. – 2023. https://doi.org/10.48550/arXiv.2304.03400.

11. Hashemi, S. H. O. Color Image Steganography using Deep Convolutional Autoencoders based on ResNet Architecture / S. H. O. Hashemi, M.-H. Majidi, & S. Khorashadizadeh // arXiv. – 2022. https://doi.org/10.48550/arXiv.2211.09409.

References

1. *HIPAA Journal. More Than 1.7 Billion Individuals Had Personal Data Compromised in 2024.* Available at: https://www.hipaajournal.com/1-7-billion-individuals-data-compromised-2024/ (accessed 18.09.2025).

2. Watts, L. *The 7 Most Telling Data Breaches of 2024. Nightfall.* – 2024. Available at: https://www.nightfall.ai/blog/the-7-most-telling-data-breaches-of-2024 (accessed 18.09.2025).

3. *Steganography. Security Vision.* Available at: https://www.securityvision.ru/education/cyberwiki/s-rus/steganografiya/ (accessed 20.09.2025) (In Russ.).

4. *Least Significant Bit (LSB). Tuple Knowledge Base.* Available at: https://www.tuple.nl/en/knowledgebase/least-significant-bit-lsb (accessed 22.09.2025).

5. *Discrete Cosine Transform (DCT). Exponenta Docs.* Available at: https://docs.exponenta.ru/images/discrete-cosine-transform.html (accessed 22.09.2025).

6. *Wavelet Transform. Gwyddion User Guide.* Available at: https://gwyddion.net/documentation/user-guide-ru/wavelet-transform.html (accessed 28.09.2025).

7. *CNSS Instructions. Committee on National Security Systems.* Available at: https://www.cnss.gov/CNSS/issuances/Instructions.cfm (accessed 28.09.2025).

8. Zhu, J., Kaplan, R., Johnson, J., & Fei-Fei, L. HiDDeN: Hiding Data With Deep Networks. *Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part XV*. Berlin, Heidelberg, Springer-Verlag, 2018, pp. 682–697. https://doi.org/10.1007/978-3-030-01267-0_40.

9. Himthani, V., Dhaka, V.S., Kaur, M., et al. Comparative performance assessment of deep learning based image steganography techniques. *Scientific Reports*, 2022, vol. 12, p. 16895. https://doi.org/10.1038/s41598-022-17362-1.

10. Bui, T., Agarwal, S., Yu, N., & Collomosse, J. (2023). RoSteALS: Robust Steganography using Autoencoder Latent Space. *arXiv*. https://doi.org/10.48550/arXiv.2304.03400.

11. Hashemi, S.H.O., Majidi, M.-H., & Khorashadizadeh, S. (2022). Color Image Steganography using Deep Convolutional Autoencoders based on ResNet Architecture. *arXiv*. https://doi.org/10.48550/arXiv.2211.09409.

Статья поступила в редакцию 28.10.2025; одобрена после рецензирования 25.11.2025; принята к публикации 25.11.2025.

The article was submitted 28.10.2025; approved after reviewing 25.11.2025; accepted for publication 25.11.2025.

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ МАШИН, КОМПЛЕКСОВ И КОМПЬЮТЕРНЫХ СЕТЕЙ

УДК 004

ИНФОРМАЦИОННАЯ СИСТЕМА ОБУЧЕНИЯ ПРОТИВОДЕЙСТВИЮ АТАКАМ МЕТОДОМ СОЦИАЛЬНОЙ ИНЖЕНЕРИИ

Кузнецова Валентина Юрьевна, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 16,

кандидат технических наук, доцент, доцент кафедры высшей и прикладной математики, ORCID 0000-0002-6954-5020, e-mail: arhelia@bk.ru

Кузнецова Екатерина Евгеньевна, Астраханский государственный университет им. В. Н. Татищева, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

старший преподаватель кафедры информационных технологий, ORCID 0009-0005-9589-0847, e-mail: katko_1994@mail.ru

Попов Алексей Викторович, Астраханский государственный университет им. В. Н. Татищева, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

студент, e-mail: alvikpopov@bk.ru

Данная статья посвящена разработке информационно-обучающей системы, предназначенной для противодействия атакам методом социальной инженерии. Рассматриваются современные угрозы информационной безопасности, связанные с человеческим фактором, подчеркивается необходимость обучения сотрудников основным методам защиты от социальной инженерии. Разработана модульная программа обучения, позволяющая последовательно формировать необходимые компетенции. Представлен детальный анализ используемых методик, включая фишинг, вишинг, смишинг и другие формы манипуляций, а также описаны особенности разработки специальной образовательной платформы, обеспечивающей практическое освоение навыков противостояния подобным атакам. Приводятся результаты экспериментов, подтверждающие повышение уровня информированности и осознанности сотрудников после прохождения соответствующего обучения. Подчеркнута важная роль регулярных тренировок и адаптации учебных материалов к современным реалиям развития угроз в области информационной безопасности

Ключевые слова: информационная безопасность, человеческий фактор, социальная инженерия, фишинг, обучение персонала, обучающая программа, тестирование, симулятор, защита данных

INFORMATION SYSTEM FOR TRAINING IN COUNTERING SOCIAL ENGINEERING ATTACKS

Kuznetsova Valentina Yu., Astrakhan State Technical University, 16 Tatishchev St., Astrakhan, 414056, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, Associate Professor of the Department of Higher and Applied Mathematics, ORCID 0000-0002-6954-5020, e-mail: arhelia@bk.ru

Kuznetsova Ekaterina E., Astrakhan Tatishchev State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

Senior Lecturer of the Department of Information Technology, ORCID 0009-0005-9589-0847, e-mail: katko_1994@mail.ru

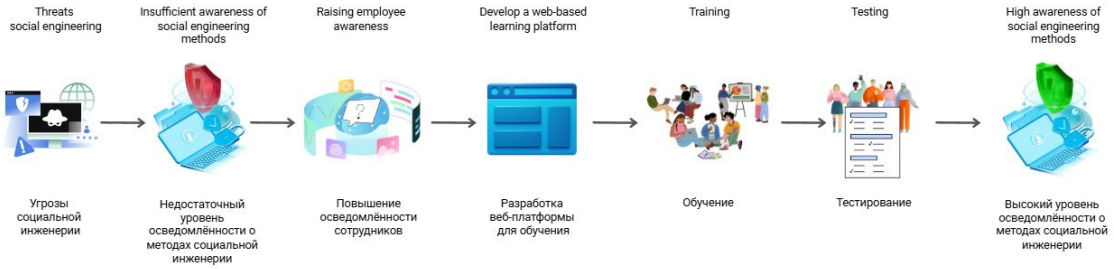
Popov Alexey V., Astrakhan Tatishchev State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

student, e-mail: alvikpopov@bk.ru

This article is devoted to the development of an information and training system designed to counter social engineering attacks. Modern threats to information security related to the human factor are considered, and the need to train employees in basic methods of protection against social engineering is emphasized. A modular training program has been developed that allows you to consistently form the necessary competencies. A detailed analysis of the techniques used, including phishing, vishing, smishing and other forms of manipulation, is presented, as well as the specifics of developing a special educational platform that provides practical skills in countering such attacks. The results of experiments are presented, confirming the increase in the level of awareness and awareness of employees after undergoing appropriate training. The important role of regular training and the adaptation of educational materials to the modern realities of the development of threats in the field of information security is emphasized.

Keywords: information security, human factor, social engineering, phishing, staff training, training program, testing, simulator, data protection

Graphical annotation (Графическая аннотация)



ВВЕДЕНИЕ

В настоящее время многие компании понимают важность безопасности данных и внедряют технические решения для защиты информации. Многие уязвимости в информационных системах закрываются различными техническими или программными методами, однако, несмотря на предпринятые меры, стабильный рост абсолютного числа проблем в области информационной безопасности сохраняется. Все из-за того, что в результате данных мер защиты, в совокупности с активным развитием технологий фокус злоумышленников сместился с техники на человека, так как он остается слабым звеном в сложных современных IT-системах. Отсюда на передний план выходит человеческий фактор в информационной безопасности – это совокупность психологических, поведенческих и социальных характеристик людей, которые могут влиять на безопасность информационных систем и данных, включая ошибки, неосторожность, невежество, предвзятости, эмоциональное принятие решений, социальное влияние и другие факторы, которые могут привести к угрозам и нарушениям безопасности [1].

Человеческий фактор в информационной безопасности компании, независимо от ее размеров и отрасли, играет ключевую роль. По данным исследований Positive Technologies, 51 % всех атак на компании в 2024 г. использует методы социальной инженерии [2]. Нарушение основных сервисов безопасности данных может привести к нарушению работы организации, финансовым и репутационным потерям, а также несоблюдению обязательств перед клиентами и партнерами. Наиболее распространенные виды уязвимостей приведены в таблице 1.

Таблица 1 – Перечень основных причин уязвимостей

Основные виды уязвимостей	Примеры
Недостаток осведомленности сотрудников	Многие сотрудники не обладают достаточными знаниями в области информационной безопасности. По статистике, только 34 % работников могут правильно определить признаки фишингового письма [2]. Они не могут распознать фишинговые письма, не понимают важности использования сложных паролей и часто не осознают рисков, связанных с использованием личных устройств для работы.
Психологические манипуляции	Злоумышленники активно используют психологические приемы, такие как создание ложного ощущения срочности («Немедленно обновите пароль!») или авторитета («Это письмо от руководителя»). Эти методы эксплуатируют естественные человеческие слабости, такие как доверие к авторитетам и страх перед наказанием.
Пренебрежение корпоративными правилами безопасности	Даже при наличии четких политик безопасности сотрудники часто их нарушают. Согласно опросу Kaspersky Lab (2023) [3], 61 % сотрудников нарушают политики безопасности, чтобы «ускорить работу»: используют простые пароли, пересылают конфиденциальные данные через незащищенные каналы или игнорируют необходимость регулярного обновления ПО. Это создает дополнительные риски для организации

Современные технические средства защиты не могут полностью компенсировать эти риски, поскольку они не устраняют основную причину – человеческое поведение. А. В. Токолов, приводя в пример слова основоположника социальной инженерии Карла Поппера, трактует определение социальной инженерии как «совокупность подходов в прикладных социальных науках, ориентированных на изменения поведения и установок людей, на разрешение социальных проблем, на адаптацию социальных институтов к изменяющимся условиям, на сохранение социальной активности» [4, с. 176]. Еще в начале XX в. компании набирали сотрудников в качестве социальных инженеров,

задача которых не только войти в доверие к потребителям, но и изменить их мнение, что в последующем привело к преступным действиям. С появлением интернета социальная инженерия стала распространяться еще быстрее. Рассылка спам-писем, поддельные сайты – все это давало возможность хакерам получить доступ защищенным данным. Согласно исследованиям А. В. Токолова [4], М. О. Янгаевой [5], И. Д. Диреева [6], В. А. Репенко, С. А. Резниченко [7], наиболее распространенные методы атак социальной инженерии включают в себя:

Фишинг (англ. phishing) остается одним из самых распространенных методов атак. Согласно исследованию Proofpoint (2023), 98 % фишинговых атак требуют активного взаимодействия пользователя – например, перехода по ссылке, открытия вложения или ввода учетных данных [8]. Это означает, что даже самые надежные системы защиты могут быть обойдены, если сотрудник не распознает угрозу. Злоумышленники рассылают поддельные письма или создают фальшивые сайты, имитирующие доверенные организации. В таких сообщениях часто используются угрозы, например, о блокировке счета или обещания больших выигрышей. Характерным признаком фишинга являются грамматические ошибки и поддельные ссылки, которые лишь внешне похожи на настоящие.

Вишинг (англ. vishing) – это телефонный вариант фишинга. Мошенники используют поддельные голосовые сообщения или звонки, представляясь сотрудниками банков или служб поддержки. Они могут просить ввести PIN-код или пароль через телефонную клавиатуру, а также создавать ложные экстренные ситуации, чтобы заставить жертву действовать быстро и необдуманно.

Смишинг (англ. smishing) – это разновидность фишинга с использованием SMS-сообщений. Его еще называют SMS-фишинг. Как и в случае с фишингом, мошенники маскируются под легитимные организации и обманом вынуждают своих жертв раскрыть конфиденциальную информацию. Для SMS-фишинга мошенники используют вредоносные программы или поддельные вебсайты. При этом может быть задействована как служба SMS, так и другие сервисы, например мессенджеры для мобильных устройств, которые позволяют обмениваться данными.

Претекстинг (англ. pretexting) основан на создании ложного предлога для получения информации. Злоумышленник тщательно готовит легенду, например представляется коллегой из IT-отдела или начальником. Используя известные факты о жертве, такие как дата рождения или номер паспорта, он усиливает доверие и добивается своей цели.

Квид про кво (англ. quid pro quo) – это метод, при котором злоумышленник предлагает что-то взамен на конфиденциальные данные. Например, он может представиться сотрудником техподдержки и предложить помощь в устранении несуществующей проблемы, попросив для этого пароль или доступ к системе.

Троянский конь (англ. trojan Horse) – это вредоносное ПО, замаскированное под полезную программу. Одной из разновидностей этого метода является «дорожное яблоко», когда злоумышленники подбрасывают зараженные флешки или диски в местах, где их могут найти сотрудники компании. Любопытство или доверие к внешнему виду носителя заставляет жертву вставить его в компьютер, что приводит к заражению системы.

Плечевой серфинг (англ. shoulder surfing) заключается в визуальном наблюдении за вводом конфиденциальной информации, такой как пароли или PIN-коды. Этот метод особенно эффективен в общественных местах, где люди менее внимательны к своему окружению.

Обратная социальная инженерия отличается тем, что жертва сама предлагает злоумышленнику нужные ему данные. Например, мошенник может создать искусственную проблему, а затем предложить ее решение, попросив у жертвы пароль или другие конфиденциальные сведения.

Защита от таких атак требует не только технических мер, но и постоянного обучения сотрудников, что является не формальностью, а ключевым элементом кибербезопасности любой организации. Методика обучения персонала противодействию атакам методами социальной инженерии для эффективного усвоения знаний и формирования устойчивых навыков должна базироваться на четырех ключевых принципах:

1) практико-ориентированность;
2) персонализация обучения;
3) непрерывность и регулярность;
4) мотивация и вовлеченность.

Так, например, в статье Д. С. Ан, А. С. Зуфарова [9] описывается важность обучения сотрудников методам защиты от социальной инженерии на примере фишинговых атак. Авторы предлагают регулярные интерактивные тренинги, симуляции и инструкции по противодействию фишингу как эффективный способ повышения уровня осведомленности. К. В. Пителинский, В. В. Бритвина, И. В. Калуцкий, С. О. Маковей, В. А. Милорадов, В. Г. Мерзликин [10] предлагают комплексный подход к обучению с использованием дистанционных обучающих платформ, подкрепленный регулярным тестированием. Все эти методы полезны для базового повышения осведомленности,

но отсутствие персонализированного подхода к обучению и интеграции с корпоративной инфраструктурой организации, развивающие и меняющиеся технологии социальной инженерии снижают уровень мотивации и практической подготовки и, как следствие, не обеспечивают достаточного вовлечения сотрудников в процесс обучения.

Современные методы обучения, такие как онлайн-курсы и вебинары, позволяют сотрудникам получать знания в удобном формате и темпе. Регулярные тренинги, симуляции атак и игровые форматы помогают закрепить полученные знания на практике. Такие подходы не только повышают уровень осведомленности о киберугрозах, но и формируют корпоративную культуру безопасности. Использование интерактивных методов обучения делает процесс более увлекательным и запоминающимся.

АНАЛИЗ ИМЕЮЩИХСЯ НА РЫНКЕ РЕШЕНИЙ И ФОРМУЛИРОВАНИЕ ТРЕБОВАНИЙ

На рынке представлены разнообразные онлайн-курсы и платформы по кибербезопасности, среди которых выделяются такие крупные игроки, как Coursera, Udemy, edX, Skillbox, Яндекс.Практикум и др. Эти курсы охватывают широкий спектр тем – от основ информационной безопасности и защиты сетей до этичного хакинга и технических аспектов безопасности:

– Coursera и edX предлагают академически выверенные программы с возможностью получения сертификатов от ведущих университетов. Их курсы хорошо подходят для глубокого теоретического изучения и подготовки специалистов, однако зачастую не фокусируются на практических аспектах социальной инженерии и психологических методах атак;

– Udemy и Skillbox предоставляют более практикоориентированные курсы с упором на технические навыки, но качество материалов может варьироваться, а внимание к человеческому фактору и социальным атакам часто ограничено;

– платформы для практического обучения (например, Hack The Box, TryHackMe) ориентированы на развитие технических навыков offensive и defensive security, включая пентестинг и работу с реальными уязвимостями. Однако они не затрагивают психологические и поведенческие аспекты, которые лежат в основе социальной инженерии;

– Cybrary и другие бесплатные ресурсы предоставляют базовые знания, но зачастую не обеспечивают глубины и системности, необходимых для формирования устойчивых навыков противодействия социальным атакам.

Проектирование специализированной веб-платформы по обучению противодействию атакам методом социальной инженерии обусловлено недостаточной эффективностью существующих решений, которые либо слишком технически ориентированы, либо не охватывают психологические и поведенческие аспекты угроз. Отсутствие интерактивных, адаптивных и практикоориентированных инструментов, направленных на развитие навыков распознавания и предотвращения социальных атак, создает пробел в образовательном процессе. Учитывая выявленные недостатки существующих решений, разработанная веб-платформа по обучению противодействию методам социальной инженерии должна отвечать следующим ключевым требованиям:

1. Фокус на человеческом факторе и социальной инженерии. В процессе обучения затрагиваются не только технические аспекты, но и психологические приемы манипуляции способом распознавания и противодействия социальным атакам, что позволяет повысить осведомленность и устойчивость обучающихся к реальным угрозам.

2. Интерактивность и практическая направленность. Важным элементом являются симуляции фишинговых атак, кейсы, тесты и игровые сценарии, которые позволяют обучающимся отрабатывать навыки в безопасной среде и получать обратную связь. Практический опыт значительно повышает эффективность обучения.

3. Адаптивность и персонализация. Платформа подстраивается под уровень знаний и потребности каждого пользователя, предлагая индивидуальные маршруты обучения и учитывая специфику отрасли и компании.

4. Регулярное обновление контента. В связи с постоянным развитием методов социальной инженерии, содержание курса своевременно обновляется, отражая новые угрозы и способы защиты.

5. Мониторинг и аналитика. Встроенные инструменты для отслеживания прогресса, выявления уязвимых зон и формирования отчетов для руководства помогают контролировать эффективность обучения и принимать управленческие решения.

6. Доступность и удобство использования. Веб-платформа доступна с различных устройств, имеет интуитивно понятный интерфейс и обеспечивает комфортное обучение в любое время.

Однако разработка самой платформы не имеет смысла без четко выверенного профессионального учебного курса противодействию атакам метода социальной инженерии. Структура такого курса должна быть рассчитана как для всех сотрудников компании, вне зависимости от их должности и уровня технической подготовки, поскольку атаки социальной инженерии могут быть направлены на любые категории персонала:

1. Рядовые сотрудники – часто становятся мишенью из-за доступа к критически важным данным и системам.

2. Руководители – представляют интерес для злоумышленников в рамках целевого фишинга (spear-phishing) из-за высокого уровня доступа и влияния.

3. IT-специалисты – несмотря на высокий уровень технической подготовки, также подвержены атакам, таким как претекстинг (создание ложной предыстории) и вишинг (фишинг по телефону).

РАЗРАБОТКА УЧЕБНОГО КУРСА

Курс предлагается реализовывать в формате онлайн-обучения на специально разработанной веб-платформе, которая учитывает модульную структуру и включает интерактивные элементы для повышения вовлеченности и усвоения материала. Общее время прохождения курса составляет 2–3 часа, при этом предусмотрена возможность разбивать обучение на несколько сессий, что обеспечивает комфортное восприятие учебного материала без перегрузки. Гибкий график позволяет сотрудникам проходить обучение в удобное время, что особенно важно для компаний с удаленными сотрудниками или нестандартным рабочим днем. Прогресс сохраняется автоматически, позволяя возвращаться к обучению с места остановки.

На основании исследований А. С. Унукович [11], C. Hadnagy [12] предложены следующие модули (табл. 2).

Таблица 2 – Модульная структура курса

Модуль	Цель	Время
Модуль 1. Введение в социальную инженерию и её угрозы	Ознакомить с понятием социальной инженерии, показать реальные случаи и последствия атак для компании и сотрудников.	20 мин
Модуль 2. Психология социальной инженерии	Рассмотреть психологические приемы и механизмы манипуляции, объяснить, почему люди становятся жертвами обмана.	25 мин
Модуль 3. Фишинг и его разновидности	Научить распознавать фишинговые письма, SMS, звонки, а также методы защиты от них.	30 мин
Модуль 4. Претекстинг и вишинг: опасные звонки и личные манипуляции	Обучить проверять подозрительные звонки и сообщения, имитирующие коллег или руководство, и правильно реагировать.	20 мин
Модуль 5. Физические угрозы и безопасность рабочего места	Предупредить о рисках использования неизвестных флешек, документов и других физических носителей информации.	15 мин
Модуль 6. Практические навыки и симуляции атак	Закрепить знания через интерактивные симуляции, кейсы и тесты, развить навыки распознавания и противодействия атакам.	30 мин
Итоговый тест	Проверить усвоение материала и способность применять знания на практике	15 мин

Пример пользовательского интерфейса информационной обучающей среды приведен на рисунке 1.

Описание блоков модуля:

1. Теоретический блок (10–15 минут).

В этом блоке слушатели получают описание конкретного типа атаки (например, фишинг, претекстинг, кви про кво), знакомятся с реальными примерами из практики и изучают психологические механизмы, которые используют злоумышленники. Также даются инструкции по распознаванию угроз и алгоритмы правильных действий.

2. Интерактивные задания (5–10 минут).

Слушатели анализируют подозрительные письма, звонки или сообщения, выбирают правильную реакцию в смоделированных ситуациях. Такой формат позволяет закрепить теоретические знания и развить практические навыки.

3. Симуляция атаки (проводится периодически).

Периодически проводится контролируемая фишинг-рассылка или имитация звонка мошенника с согласия руководства компании. После теста проводится разбор ошибок с тренером или автоматизированная обратная связь, что помогает выявить слабые места и скорректировать поведение.

4. Разбор ошибок.

Если сотрудник «попался» на уловку, он получает персонализированные рекомендации, например: «Вы перешли по ссылке в письме. Обратите внимание на домен – в данном случае он не соответствовал официальному». Это способствует закреплению правильных моделей поведения.

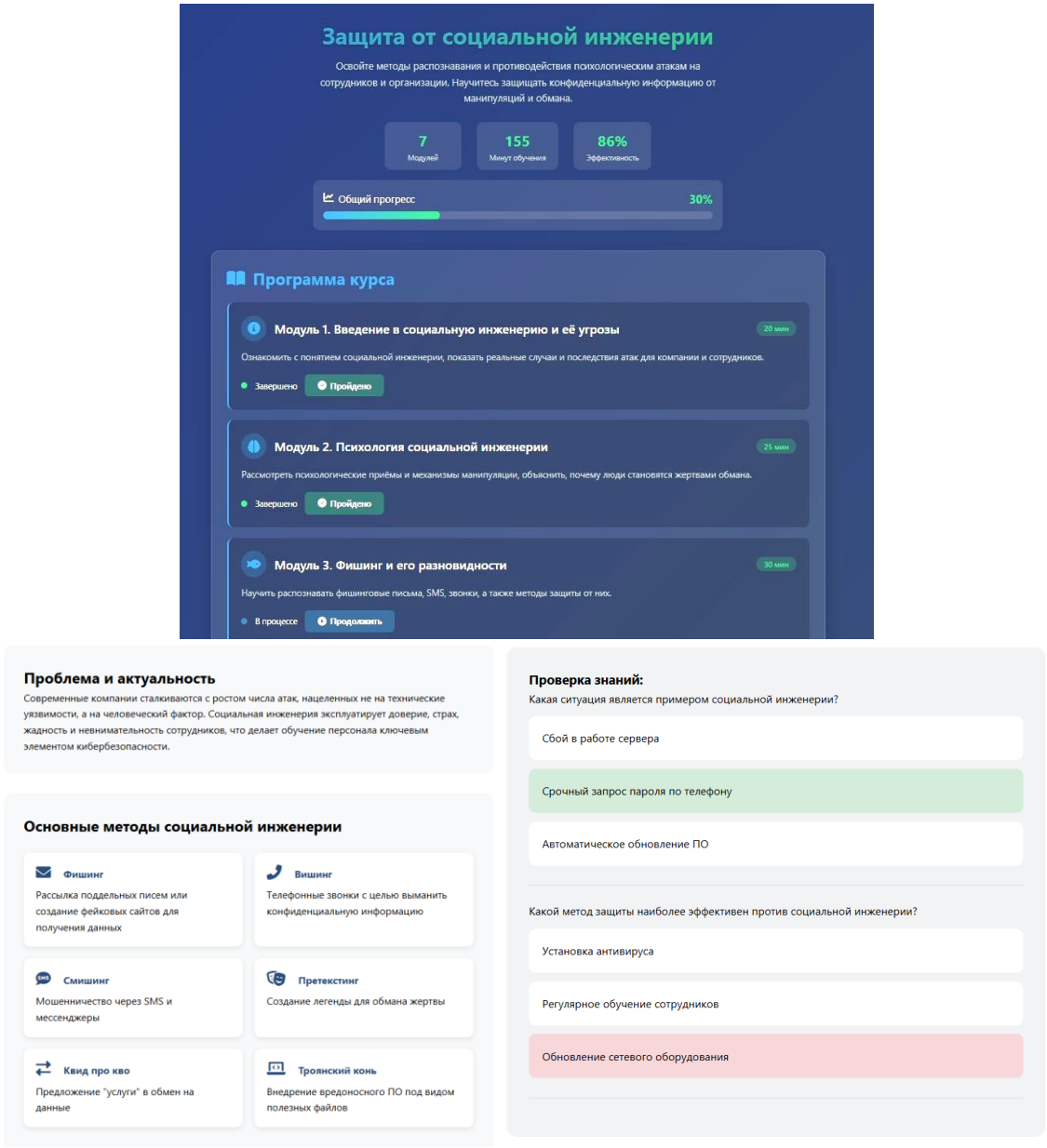


Рисунок 1 – Теоретическая (слева) и тестовая (справа) части обучающей платформы

Отдельно авторами делается акцент на симуляции различных видов атак, где обучающийся в режиме реального времени может погрузиться в реальную ситуацию, будь то получение фишингового письма или звонок от мошенника. При этом система предлагает в случайном порядке как безопасные ситуации, так и мошеннические кейсы (рис. 2).

Исходя из анализа современных курсов и рекомендаций экспертов, структура курса сбалансирована и охватывает ключевые аспекты социальной инженерии. При необходимости можно дополнить курс модулем по развитию корпоративной культуры безопасности и мотивации сотрудников, а также разделом по работе с инцидентами и алгоритмам реагирования при подозрении на атаку. Однако для базового курса, рассчитанного на широкий круг пользователей, текущая структура оптимальна, она обеспечивает постепенное погружение в тему и сочетает теорию с практикой.

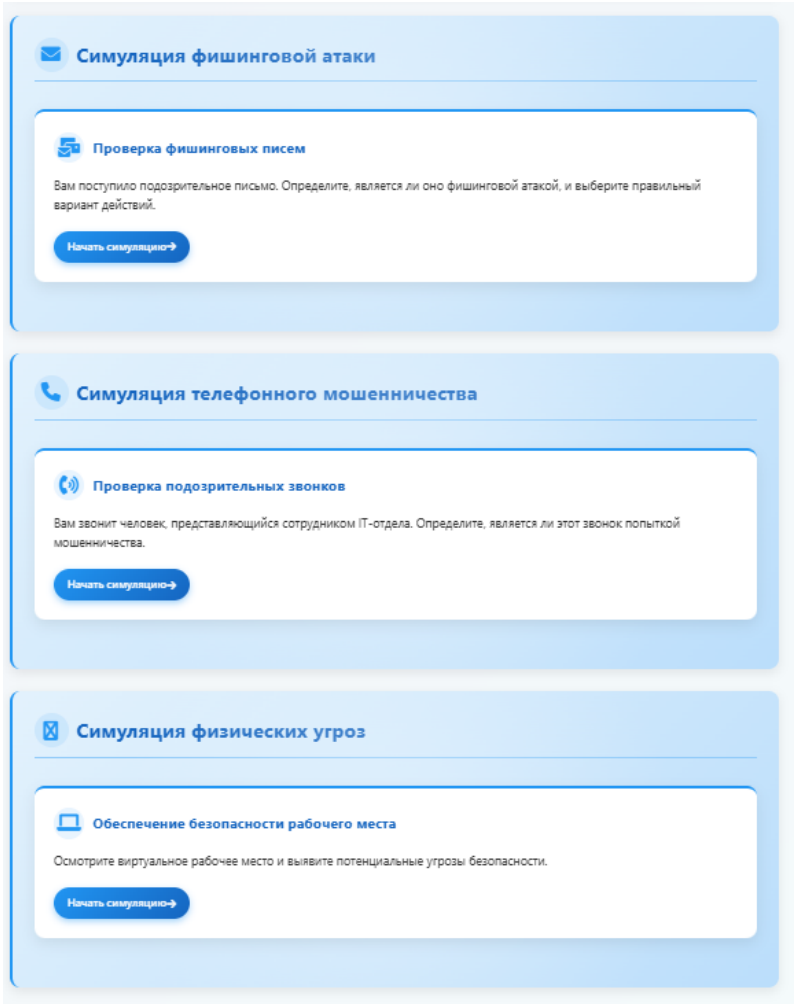


Рисунок 2 – Пример кейсов

Что касается реализации самой платформы, backend-часть веб-платформы была разработана на стыке технологий SQLAlchemy, Pydantic и FastAPI, что обеспечивает высокую производительность, надежность и защиту от SQL-иньекций. Пользовательский интерфейс реализован с помощью базовых функций HTML и CSS.

Апробация и результаты веб-платформы проводились с помощью экспериментальной и контрольной групп из числа линейного персонала государственных и коммерческих организаций города Астрахани. Всего в исследовании приняли участие 113 сотрудников из различных подразделений, которых случайным образом разделили на две равные группы:

– экспериментальная группа – сотрудники, прошедшие инструктаж и обучение на онлайн-платформе (57 человек);

– контрольная группа – сотрудники, которые не проходили дополнительное обучение и продолжали работать в обычном режиме без специальных тренингов (60 человек).

Обе группы прошли предварительное тестирование для оценки исходного уровня знаний и навыков в области защиты от социальной инженерии. Результаты предварительного тестирования представлены на рисунке 3.

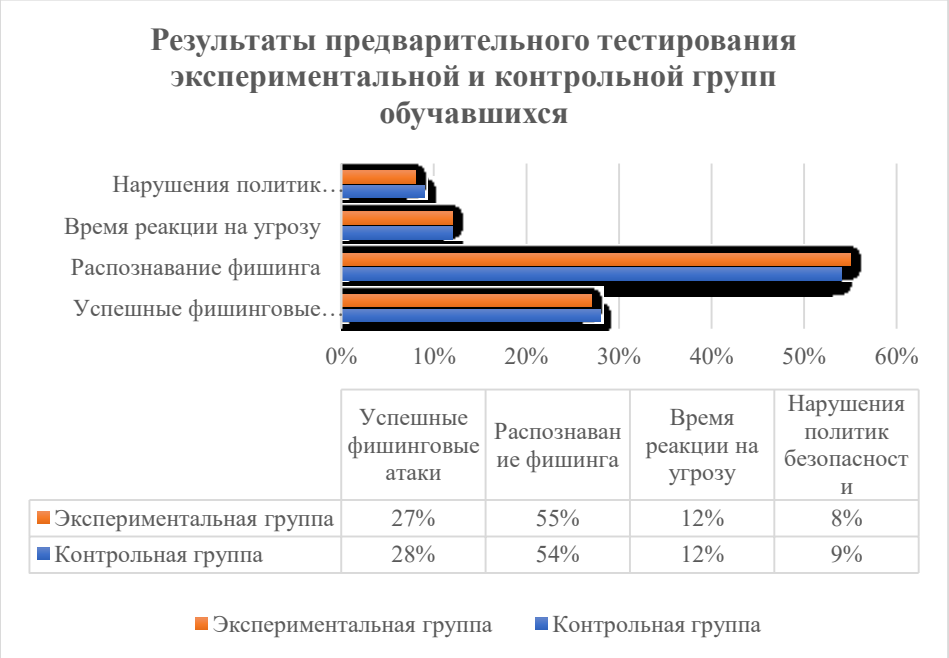


Рисунок 3 – Результаты предварительного тестирования экспериментальной и контрольной групп обучавшихся

После того как экспериментальная группа прошла обучение на платформе, при условии, что контрольная группа проходила обучение самостоятельно с помощью открытых платформ, при этом все участники эксперимента прошли итоговое тестирование. Анализировались показатели успешности распознавания угроз, время реакции, количество политик безопасности и уровень вовлеченности в обучение (для экспериментальной группы). В результате эксперимента экспериментальная группа показала значительные улучшения ключевых показателей безопасности по сравнению с контрольной группой (результаты эксперимента приведены на рисунке 4).

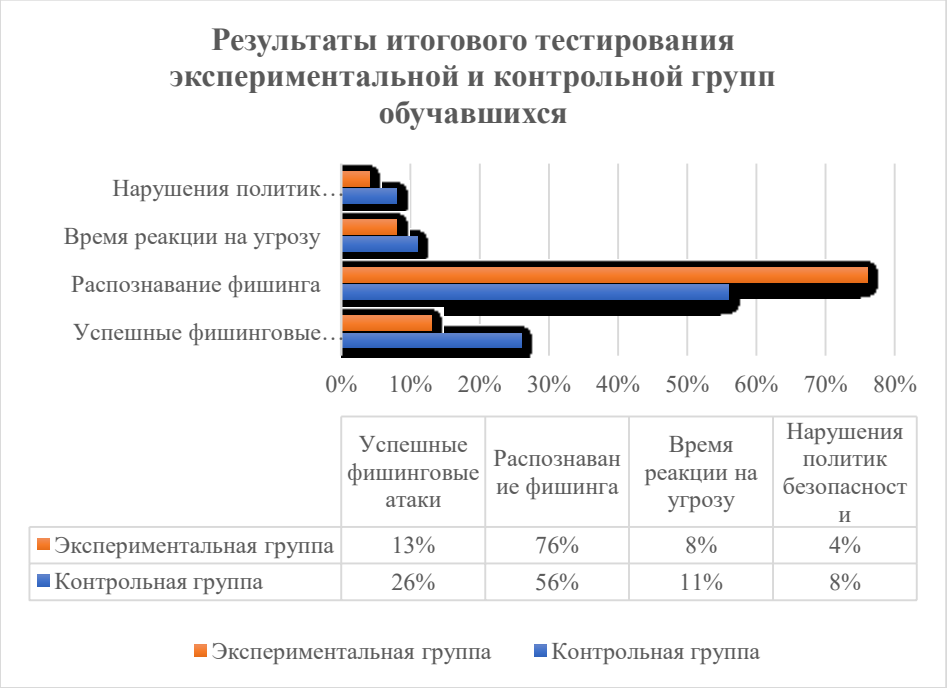


Рисунок 4 – Результаты итогового тестирования экспериментальной и контрольной групп обучавшихся

Экспериментальная группа показала в 1,35 раз выше эффективность освоения знаний по сравнению с контрольной группой. Таким образом, внедрение веб-платформы с модульной системой обучения и симуляциями реальных атак повысило устойчивость к социальной инженерии на 35 % по сравнению с типовыми решениями.

ЗАКЛЮЧЕНИЕ

В статье рассмотрена задача обеспечения безопасности данных, сделан акцент на антропогенных угрозах, ключевой из которых являются методы социальной инженерии. Авторами была предложена веб-платформа, которая позволяет обучить пользователей основным методам защиты от указанного вида угроз в рамках комбинированного модульного обучения с использованием различных видов активностей. Была собрана экспериментальная группа и проведена апробация, результат которой показал повышение устойчивости к атакам.

Список источников

1. Лебедев, М. А. Роль человеческого фактора в информационной безопасности / М. А. Лебедев, Д. А. Демкин // Вестник науки. – 2024. – № 8 (77). – С. 141–145.

2. Аналитическая статья Positive Technologies, Актуальные киберугрозы: IV квартал 2024 года – I квартал 2025 года. – URL: https://ptsecurity.com/ru-ru/research/analytics/aktualnye-kiberugrozy-iv-kvartal-2024-goda-i-kvartal-2025-goda/#id1 (дата обращения 10.05.2025).

3. Kids Safe Media. Взрослые и дети в интернете: аналитический отчет 2023. – URL: https://kids.kaspersky.ru/article/vzroslye_i_deti_v_internete_analiticheskiy_otchet_2023 (дата обращения 10.05.2025).

4. Токолов, А. В. Социальная инженерия в вопросах обеспечения информационной безопасности / А. В. Токолов // Криминологический журнал. – 2024. – № 4. – С. 175–182.

5. Янгаева, М. О. Социальная инженерия как способ совершения киберпреступлений / М. О. Янгаева // Вестник Сибирского юридического института МВД России. – 2021. – № 1 (42). – С. 133–138.

6. Диреев, И. Д. Социальная инженерия в контексте информационной безопасности / И. Д. Диреев // МНИЖ. – 2025. – № 3 (153). – С. 1–8.

7. Репенко, В. А. Защита от атак с применением средств и методов социальной инженерии / В. А. Репенко, С. А. Резниченко // Вестник ДГТУ. Технические науки. – 2022. – № 4. – С. 85–96.

8. Proofpoint Research, 2023 Ponemon Healthcare Cybersecurity Report. – URL: https://www.proofpoint.com/uk/resources/threat-reports/ponemon-healthcare-cybersecurity-report (дата обращения 10.05.2025).

9. Ан, Д. С. Влияние интерактивных тренингов на осведомленность сотрудников о фишинговых атаках / Д. С. Ан, А. С. Зуфарова // ЦИТИСЭ. – 2025. – № 1 (43). – С. 41–52.

10. Пителинский, К. В. О некоторых решениях, повышающих осведомленность персонала компании в противодействии социальной инженерии / К. В. Пителинский, В. В. Бритвина, И. В. Калуцкий, С. О. Маковей // Информационная безопасность. – 2024. – № 1. – С. 46–53.

11. Унукович, А. С. Социальная инженерия и кибербезопасноть: виктимологический аспект / А. С. Унукович // Психопедагогика в правоохранительных органах. – 2021. – № 3 (86). – С. 346–351.

12. Hadnagy, C. Social Engineering: The Science of Human Hacking / C. Hadnagy. – 2nd ed. – Hoboken : John Wiley & Sons, 2018. – 368 p.

References

1. Lebedev, M. A., Demkin, D. A. The Role of the Human Factor in Information Security. *Science Bulletin*, 2024, no. 8 (77), pp. 141–145 (In Russ.).

2. *Analytical Article by Positive Technologies, Current Cyber Threats: Q4 2024 – Q1 2025.* Available at: https://ptsecurity.com/ru-ru/research/analytics/aktualnye-kiberugrozy-iv-kvartal-2024-goda-i-kvartal-2025-goda/#id1 (accessed 10.05.2025) (In Russ.).

3. *Kids Safe Media. Adults and Children on the Internet: Analytical Report 2023.* Available at: https://kids.kaspersky.ru/article/vzroslye_i_deti_v_internete_analiticheskiy_otchet_2023 (accessed 10.05.2025) (In Russ.).

4. Tokolov, A. V. Social Engineering in Information Security Issues. *Criminological Journal*, 2024, no. 4, pp. 175–182 (In Russ.).

5. Yangaeva, M. O. Social Engineering as a Method of Cybercrime. *Bulletin of the Siberian Law Institute of the Ministry of Internal Affairs of Russia*, 2021, no. 1 (42), pp. 133–138 (In Russ.).

6. Direev, I. D. Social Engineering in the Context of Information Security. *MNIZH*, 2025, no. 3 (153), pp. 1–8 (In Russ.).

7. Repenko, V. A., Reznichenko, S. A. Protection Against Attacks Using Social Engineering Tools and Methods. *Bulletin of Dagestan State Technical University. Technical Sciences*, 2022, no. 4, pp. 85–96 (In Russ.).

8. *Proofpoint Research, 2023 Ponemon Healthcare Cybersecurity Report.* Available at: https://www.proofpoint.com /uk/resources/threat-reports/ponemon-healthcare-cybersecurity-report (accessed 10.05.2025).

9. An, D. S., Zufarova, A. S. The influence of interactive training on employees' awareness of phishing attacks. *CITISE*, 2025, no. 1 (43), pp. 41–52 (In Russ.).

10. Pitelinsky, K. V., Britvina, V. V., Kalutsky, I. V., Makovey, S. O. On some solutions for increasing the awareness of company personnel in countering social engineering. *Information Security*, 2024, no. 1, pp. 46–53 (In Russ.).

11. Unukovich, A. S. Social Engineering and Cybersecurity: A Victimological Aspect. *Psychopedagogy in Law Enforcement Agencies*, 2021, no. 3 (86), pp. 346–351 (In Russ.).

12. Hadnagy, C. *Social Engineering: The Science of Human Hacking*. 2nd ed. Hoboken, John Wiley & Sons, 2018. 368 p.

Статья поступила в редакцию 15.10.2025; одобрена после рецензирования 17.11.2025; принята к публикации 21.11.2025.

The article was submitted 15.10.2025; approved after reviewing 17.11.2025; accepted for publication 21.11.2025.

УДК 004.85

РАЗРАБОТКА МЕТОДА ПРОЕКТИРОВАНИЯ СИМУЛЯТОРОВ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ ДЛЯ ПОДГОТОВКИ СТУДЕНТОВ ИНЖЕНЕРНЫХ СПЕЦИАЛЬНОСТЕЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ

Асланов Роман Эдвинович, ГБПОУ г. Москвы «Московский государственный образовательный комплекс», 115114, Российская Федерация, г. Москва, 2-й Павелецкий пр-д, 4А,

кандидат технических наук, начальник отдела информационных технологий, e-mail: aslanov.boxing@mail.ru

Большаков Александр Афанасьевич, Санкт-Петербургский политехнический университет Петра Великого, 195251, Российская Федерация, г. Санкт-Петербург, ул. Политехническая, 29,

доктор технических наук, профессор, e-mail: aabolshakov57@gmail.ru

Вешнева Ирина Владимировна, Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, 410012, Российская Федерация, г. Саратов, ул. Астраханская, 83,

доктор технических наук, доцент, профессор, e-mail: veshnevaiv@mail.ru

Ключиков Аркадий Викторович, Вавиловский университет, 410012, Российская Федерация, г. Саратов, пр-т им. Петра Столыпина, зд. 4, стр. 3,

кандидат технических наук, заведующий кафедрой, e-mail: krok9407@mail.ru

Малько Евгений Игоревич, ООО «Лаборатория Наносемантика», 121 357, Российская Федерация, г. Москва, ул. Верейская, д. 29, стр. 134,

Middle Unity разработчик отдела 3D и GameDev, e-mail: i@malko-pro.ru

Предложен современный метод проектирования симуляторов виртуальной реальности для инженерного образования, основанный на онтологическом моделировании. Исследование направлено на решение проблемы фрагментарности существующих подходов к разработке VR-тренажеров, обусловленной отсутствием стандартизированных методов проектирования. В работе детально описана разработанная онтологическая модель предметной области, формализующая взаимосвязи между профессиональными компетенциями, учебными материалами и педагогическими сценариями. Предложен структурированный метод проектирования, включающий три ключевых этапа с детализацией процессов концептуализации, итеративной разработки и тестирования. Особое внимание уделено методологии формирования исходных видеоданных производственных процессов. Апробация метода выполнена на примере создания симуляторов для подготовки операторов металлорежущих станков, подтвердившая эффективность предложенного подхода для создания сложных технических обучающих систем. Использование предложенных решений позволит существенно повысить качество образовательных VR-продуктов на основе обеспечения их соответствия требованиям профессиональных стандартов и образовательных программ.

Ключевые слова: виртуальная реальность, симулятор, компьютерный тренажер, онтологическая модель

Финансирование: *исследование выполнено при поддержке Российского научного фонда, грант № 25-21-00334, «Прототип системы анализа психофизического состояния человека при обучении методами иммерсивных технологий», https://rscf.ru/project/25-21-00334/, выполняемый в Санкт-Петербургском политехническом университете Петра Великого.*

DEVELOPMENT OF A METHOD FOR DESIGNING VIRTUAL REALITY SIMULATORS FOR TRAINING ENGINEERING STUDENTS BASED ON AN ONTOLOGICAL MODEL

Aslanov Roman E., SBPEI of Moscow "Moscow State Educational Complex", 4A 2nd Paveletsky Proezd, Moscow, 115114, Russian Federation,

Cand. Sci. (Engineering), Head of the Information Technology Department, e-mail: aslanov.boxing @mail.ru

Bolshakov Aleksandr A., Peter the Great St. Petersburg Polytechnic University, 29 Politekhnicheskaya St., St. Petersburg, 195251, Russian Federation,

Doct. Sci., (Engineering), Professor, e-mail: aabolshakov57@gmail.ru

Veshneva Irina V., Saratov State University, 83 Astrakhanskaya St., Saratov, 410012, Russian Federation,

Doct. Sci., (Engineering), Associate Professor, Professor, e-mail: veshnevaiv@mail.ru

Klyuchikov Arkady V., Vavilov University, 4, Bldg. 3 Pyotr Stolypin Ave., Saratov, 410012, Russian Federation,

Cand. Sci. (Engineering), Head of Department, e-mail: krok9407@mail.ru

Malko Evgeny I., Nanosemantics Laboratory LLC, 29 Vereyskaya St., Building 134, Moscow, 121357, Russian Federation

Middle Unity Developer of 3D and GameDev Department, e-mail: i@malko-pro.ru

A modern method for designing virtual reality simulators for engineering education based on ontological modeling is proposed. This study aims to address the fragmentation of existing approaches to developing VR simulators, which is due to the lack of standardized design methods. The paper describes in detail the developed ontological model of the subject area, formalizing the relationships between professional competencies, educational materials, and pedagogical scenarios. A structured design method is proposed, including three key stages detailing the processes of conceptualization, iterative development, and testing. Particular attention is paid to the methodology for generating initial video data of production processes. The method was tested using simulators for training machine tool operators, confirming the effectiveness of the proposed approach for creating complex technical training systems. The use of the proposed solutions will significantly improve the quality of educational VR products by ensuring their compliance with the requirements of professional standards and educational programs.

Keywords: virtual reality, simulator, computer trainer, ontological model

Financial support: *the study was supported by the Russian Science Foundation, grant № 25-21-00334, "Prototype of a system for analyzing the psychophysical state of a person during training using immersive technologies," https://rscf.ru/project/25-21-00334/, carried out at Peter the Great St. Petersburg Polytechnic University.*

Графическая аннотация (Graphical annotation)

Оценка характеристик VR-симуляторов, выбор для реализации подготовки студентов инженерных специальностей

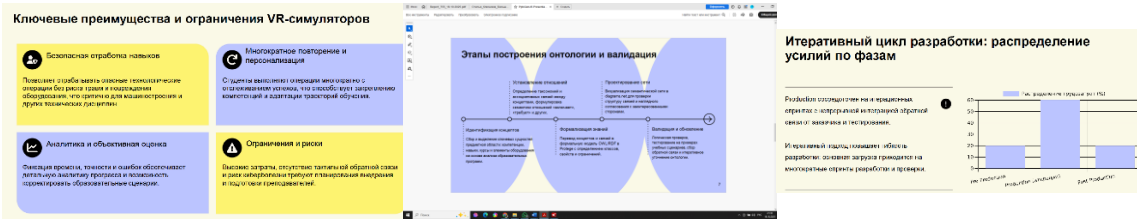
(Evaluation of VR simulators' characteristics and their selection for training engineering students)

Построение онтологической модели заданной предметной области для подготовки студентов инженерных специальностей

(Construction of an ontological model of a given subject area for training students in engineering specialties)

Этапы метода проектирования симуляторов виртуальной реальности для подготовки студентов инженерных специальностей

(Steps in the design method of virtual reality simulators for training engineering students)



ВВЕДЕНИЕ

Современное инженерное образование связано с необходимостью формирования у студентов практических компетенций в условиях цифровой трансформации производственных процессов. Традиционные методы обучения демонстрируют ограниченную эффективность при освоении сложных технологических операций и работе на специализированном оборудовании. В этой связи особую актуальность приобретают иммерсивные технологии, в частности с использованием симуляторов виртуальной реальности (VR) [11, 16], способных создавать безопасные и контролируемые иммерсивные среды для отработки профессиональных навыков [17, 23].

Одним из ключевых преимуществ VR-симуляторов является возможность создания безопасной учебной среды для отработки практических навыков. Студенты инженерных специальностей могут осваивать работу со сложным и потенциально опасным оборудованием без риска получения травм или порчи дорогостоящих производственных образцов. Это особенно актуально для таких дисциплин, как машиностроение, электротехника или химическая технология, в которых ошибки на реальном производстве могут привести к серьезным последствиям [6, 9, 21]. При этом VR-симуляторы позволяют многократно повторять технологические операции до полного их освоения, формируя устойчивые профессиональные компетенции. Существенным достоинством является также высокая степень наглядности и интерактивности обучения. Иммерсивная среда обеспечивает глубокое погружение в изучаемые процессы, позволяя визуализировать сложные физические явления и внутренние механизмы работы оборудования [3, 24]. Студенты могут наблюдать процессы, недоступные для непосредственного восприятия в реальных условиях, например течение жидкостей в гидравлических системах или распределение напряжений в конструкционных материалах. Это способствует более полному пониманию фундаментальных принципов инженерных дисциплин. К числу преимуществ следует отнести и возможность объективной оценки результатов обучения. VR-симуляторы позволяют отслеживать и анализировать каждое действие студента, фиксируя время выполнения операций, точность соблюдения технологических последовательностей и количество допущенных

ошибок. Это обеспечивает персонализированный подход к обучению и позволяет своевременно корректировать образовательную траекторию. Преподаватель получает детальную аналитику по каждому обучающемуся, что особенно ценно при работе с большими группами студентов [4, 12, 15].

Между тем использование VR-симуляторов имеет и определенные ограничения. Наиболее существенным недостатком является высокая стоимость разработки и внедрения качественных обучающих систем. Создание реалистичных виртуальных сред требует значительных финансовых вложений в специализированное оборудование, программное обеспечение и привлечение квалифицированных разработчиков. Это может стать серьезным препятствием для широкого внедрения VR-технологий в образовательный процесс, особенно в учреждениях с ограниченным бюджетом. Еще одним существенным недостатком является техническая и методологическая сложность интеграции VR-симуляторов в учебный процесс. Причем наиболее совершенные виртуальные среды не могут полностью заменить реальный практический опыт, особенно в аспекте тактильных ощущений и работы с физическими материалами. Кроме этого, существуют риски возникновения киберболезни при длительном использовании VR-оборудования, а также необходимость специальной подготовки преподавательского состава. Эти факторы требуют тщательного планирования и поэтапного внедрения VR-технологий в инженерное образование [1, 8, 10].

Процесс проектирования таких симуляторов характеризуется значительной сложностью, обусловленной необходимостью интеграции знаний из различных предметных областей: инженерных дисциплин, педагогического дизайна и компьютерных технологий. Отсутствие стандартизированных методов проектирования приводит к созданию разрозненных решений, слабо ориентированных на конкретные образовательные результаты и требования профессиональных стандартов [8, 14].

Существующие подходы к разработке VR-симуляторов часто ограничиваются технической реализацией без должного методологического обоснования. Это проявляется в несогласованности компонентов обучающей системы, недостаточной проработанности педагогических сценариев и сложности адаптации под изменяющиеся требования образовательных программ [13, 19, 20, 22]. Перспективным направлением решения указанных проблем является применение онтологического моделирования, позволяющего формализовать знания предметной области и установить семантические связи между элементами образовательного процесса. Онтологическая модель может служить концептуальным каркасом для проектирования симуляторов, обеспечивая целостность и согласованность их компонентов.

Целью работы является построение метода проектирования симуляторов виртуальной реальности для подготовки студентов инженерных специальностей на основе онтологической модели. Для ее достижения требуется решить ряд задач: осуществить анализ исследуемой предметной области и выделить ключевые сущности процесса инженерной подготовки, разработать онтологическую модель проектирования VR-симуляторов для инженерного образования, определить структуру и компоненты метода проектирования, апробировать разработанный метод на примере создания симулятора для обучения работе на металлорежущем оборудовании.

Научная новизна исследования заключается в разработке целостного метода проектирования VR-симуляторов, интегрирующего онтологическое моделирование предметной области [7], принципы педагогического дизайна и требования профессиональных стандартов. Практическая значимость работы состоит в создании методического инструментария, позволяющего систематизировать процесс разработки образовательных VR-симуляторов и обеспечить их соответствие целевым показателям качества подготовки инженерных кадров. Предложенный метод прошел апробацию в рамках разработки симуляторов для подготовки студентов по направлениям «Технология машиностроения» и «Автоматизация технологических процессов», продемонстрировав эффективность для создания сложных технических обучающих систем.

ПОСТРОЕНИЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ

Онтологическая модель определяется как кортеж:

$$OntMod =< Concept, Semantic, Interpretation >,$$

где $Concept = \{C_1, C_2, \dots, C_m\}$ – совокупность понятий, которые принадлежат исследуемой предметной области подготовки студентов инженерных специальностей;

$Semantic = \{S_1, S_2, \dots, S_n\}$, где $Semantic \subseteq \{C_1 \times C_2 \times \dots \times C_n\}$ – совокупность выявленных семантических отношений между понятиями множества $Concept$;

$Interpretation = \{ Concept \times Semantic\}$ – совокупность интерпретирующих зависимостей, которые задаются на множестве отношений. Причем интерпретирующие зависимости обычно задаются глоссарием на совокупности $Concept$.

Для представленной модели в семантической сети каждый узел связан с заданным понятием, например «знание», «технология обучения», а соединяющие их дуги соответствуют определенным типам отношения, например «требует», «использует».

Разработка онтологической модели обычно содержит определенные шаги: выявление основных понятий, определенных взаимосвязей, их формализация, построение получаемой онтологии, ее верификация и адаптация. Для формализации знаний, связанных с подготовкой студентов инженерных специальностей, была разработана онтологическая модель, представленная в виде семантической сети на рисунке 1. Онтологическая модель инженерного образования представляет целостную систему, в которой ядром является учебный процесс с его основными участниками – студентом и преподавателем. На основе учебных курсов и дисциплин формируются профессиональные компетенции, которые включают конкретные навыки и требуют соответствующих знаний.

Процесс обучения основан на ключевых педагогических принципах: активном обучении, практико-ориентированности и персонализации. Технологическая составляющая реализуется с использованием иммерсивных технологий: VR-тренажеры, мультимедийные комплексы и симуляторы.

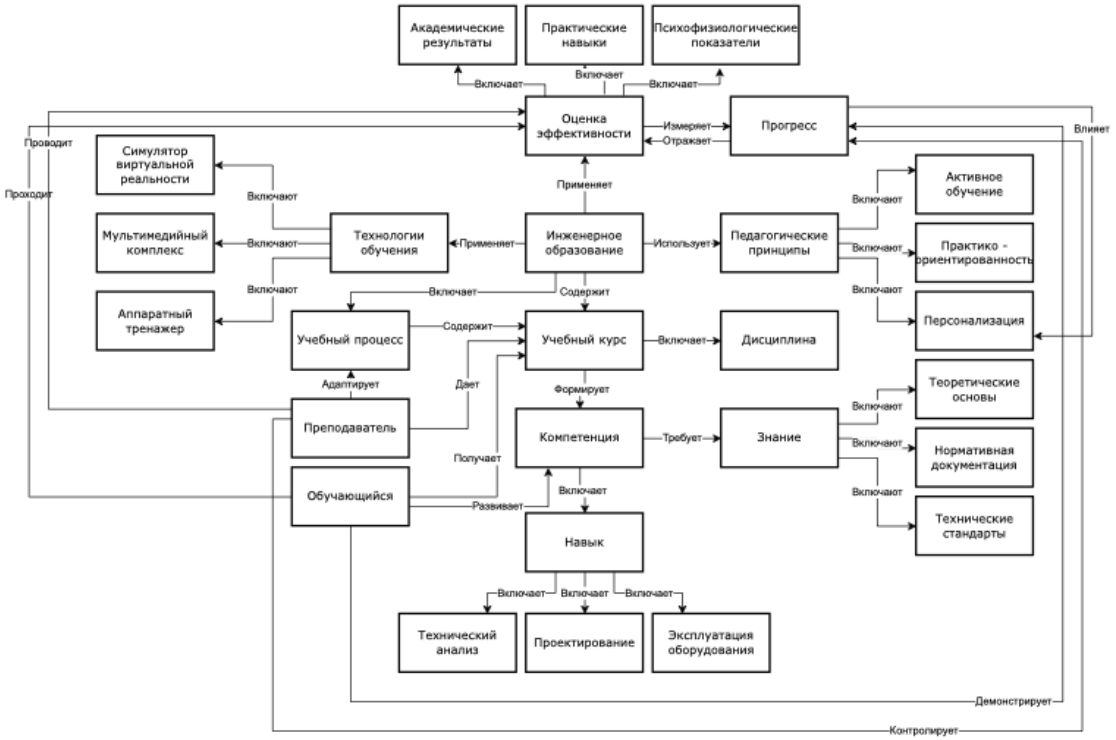


Рисунок 1 – Семантическая сеть предметной области инженерного образования

Система оценки эффективности образования является многоуровневой и включает отслеживание прогресса студента через академические результаты, практические навыки и психофизиологические показатели. Преподаватель анализирует этот прогресс и при необходимости адаптирует учебный процесс, обеспечивая непрерывную обратную связь и персонализацию обучения.

Модель устанавливает четкие семантические связи между всеми элементами, создавая основу для построения адаптивных образовательных траекторий и интеллектуальных систем управления обучением.

В качестве предметной области для апробации методологии рассмотрена деятельность операторов металлорежущих станков.

Построение семантической сети, представленной на рисунке 2, осуществлялось на основе инструментального программного средства https://app.diagrams.net/. Причем формальная реализация онтологии выполнена в среде Protege – открытом программном обеспечении для создания и управления онтологиями, разработанном Stanford Center for Biomedical Informatics Research. Его выбор обусловлен широкими возможностями по редактированию классов, свойств и индивидов, поддержкой стандартов семантического веба (OWL, RDF), наличием встроенных средств логической валидации и синтаксического контроля, а также способностью к интеграции с внешними базами данных. Визуализация разработанной онтологии приведена на рисунке 3 [2, 5].

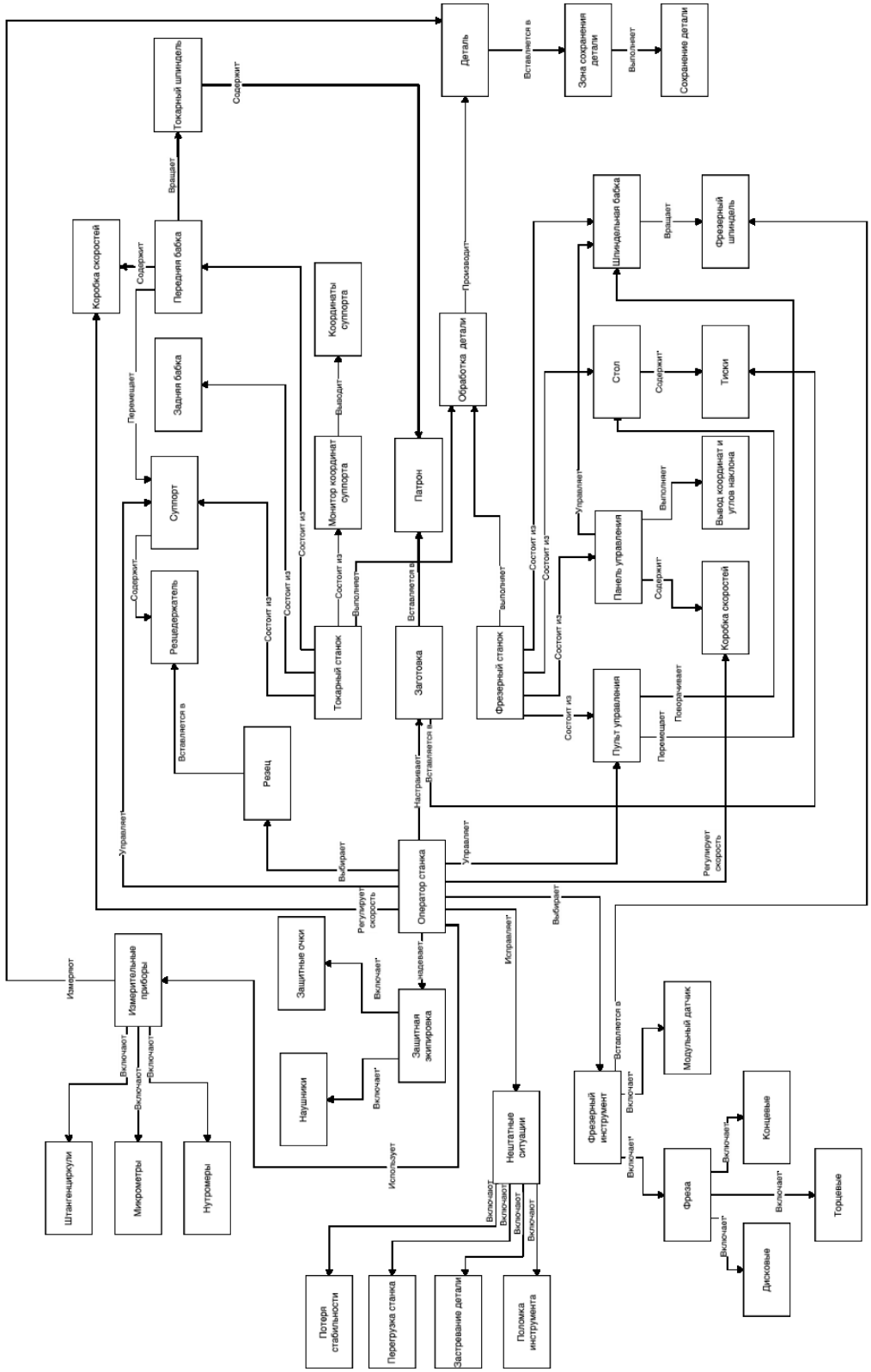


Рисунок 2 – Семантическая сеть предметной области подготовки студентов инженерных специальностей

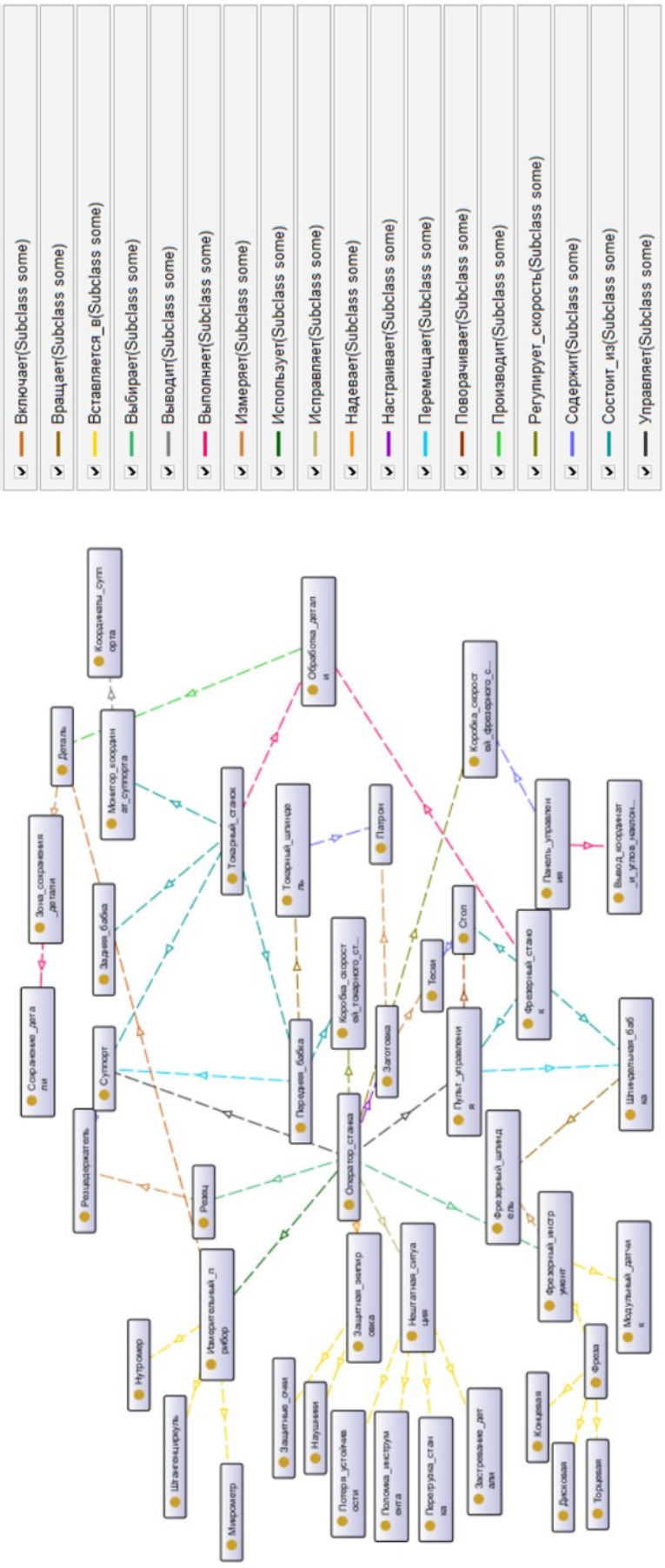


Рисунок 3 – Онтология деятельности операторов металлорежущих станков как примера предметной области

ПОСТРОЕНИЕ МЕТОДА ПРОЕКТИРОВАНИЯ СИМУЛЯТОРОВ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ ДЛЯ ПОДГОТОВКИ СТУДЕНТОВ ИНЖЕНЕРНЫХ СПЕЦИАЛЬНОСТЕЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ

Предложенный метод заключается в системной формализации жизненного цикла разработки, а также во внедрении специализированного этапа взаимодействия с заказчиком. Ключевым элементом данного этапа является использование в качестве базового артефакта функциональных видеороликов, полученных в условиях реального производственного процесса. Метод формирования исходных видеоданных регламентирует процедуру съемки на производственном объекте и включает следующие требования: видеоматериалы должны характеризовать технологический объект с точки зрения различных стадий функционирования, включая процедуру его запуска, дальнейшей работы для реализации предусмотренного технологического регламента производства требуемой продукции. Основной аспект, который должен выполняться при этом, – это видеофиксация поведения оператора для реализации заданных технологических операций при работе на аппаратуре. Для получения требуемой видеоинформации необходимо обеспечить видеосъемку, чтобы зафиксировать выполнение заданных технологических операций с различных точек, с разным масштабированием, с соответствующим близким к реальным звуковым сопровождением для наиболее полного создания производственной атмосферы при создании соответствующей иммерсивной технологии управления аппаратурой.

Для этого формируются требования к характеристикам видеофайла, которые включают следующие параметры: длительность видеофайла, степень разрешения видеозаписи, частота снимаемых кадров. Обычно рекомендуется формат видеозаписи MP4. Следует отметить, что видеофайлы снимаются по установленному регламенту, который отражает основные этапы функционирования технологического оборудования, начиная от его запуска, далее – основные режимы функционирования, а также генерация аварийных ситуаций для отработки действий по их парированию.

Ниже приводится теоретико-множественная модель, описывающая вышеуказанные факторы:

$$V_{vidproc} = f\left(T_{vidfile}, P_{frameres}, V_{descript}, N_{scenes}, C_{pressound}, Fd_{descrformat}, Ps_{prepsc}\right), \quad (1)$$

где V_{vidproc} – видеофайл с описанием технологического процесса D_{vidfile} – продолжительность видеофайла;

P_{frameres} – характеристика кадров видеофайла: частота, а также степень разрешения;

V_{descript} – контент видеофайла, содержащий описание основных режимов функционирования технологической аппаратуры;

N_{scenes} – количество видов, снимаемых сценарием из различных позиций;

$C_{\text{pressound}}$ – наличие звукового сопровождения;

$Fd_{\text{descrformat}}$ – описание видеоформата;

Ps_{prepsc} – подготовленный сценарий для построения видеофайла.

Предлагаемая процедура метода показана на рисунке 4 в виде соответствующей когнитивной схемы. Предложенная схема включает 3 блока, которые описывают основные этапы построения симуляторов иммерсивных тренажеров для предметной области подготовки студентов инженерных специальностей.

Современная практика создания виртуальных тренажеров для инженерного образования требует системного подхода, обеспечивающего управляемость и предсказуемость жизненного цикла разработки. Процесс создания VR-симулятора структурируется как последовательность взаимосвязанных этапов, каждый из которых вносит детерминированный вклад в конечный результат.

Исходная фаза проектирования охватывает комплекс процессов концептуализации и планирования. На этом этапе осуществляется анализ производственных материалов и требований заказчика, формируется техническое видение продукта, производится оценка ресурсных потребностей и составляется проектная документация. Особое внимание уделяется формированию команды разработчиков с учетом необходимых компетенций и распределения функциональных ролей. Также разрабатывается детальный план-график выполнения работ с установлением ключевых контрольных точек:

$$Pr_{pr} = f\left(Develop_{\text{concept}}, Plan_{budget}, Develop_{project}, Group_{specialists}, Implementation_{project}\right),$$

где Pr_{pr} – результаты предпроектного обследования;

$Develop_{\text{concept}}$ – построение основных концептуальных положений создания проекта с использованием видеофайлов функционирования технологического оборудования и работы операторов;

$Plan_{budget}$ – анализ и планирование бюджетных расходов;

$Develop_{project}$ – разработка Устава проекта;

$Group_{specialists}$ – набор группы специалистов-разработчиков для реализации проекта;

$Implementation_{project}$ – подготовка плана выполнения проекта.

Базовый этап построения и реализации проекта осуществляется на основе итерационного подхода. При этом весь процесс подразделяется на определенную совокупность спринтов. Каждая итерация предполагает выполнение базовых операций, которые включают идентификацию и планирование требующих решения задач, собственно, их выполнение, анализ полученных результатов с учетом мнения заинтересованных сторон. Это соответствует гибкой методологии Agile, позволяющей реализовать оперативное изменение реализуемого проекта требованиям заинтересованных сторон с учетом определенных ограничений на ресурсы и целевых показателей эффективности:

$$Project_{phase} = \sum_{i=1}^{Ns} Iteration_i\,(Ordered_i, Resplan_i, Formteam_i, Results_i, Feedback_{cust}), \quad (3)$$

где $Iteration_i$ – i-й итерационный спринт проектного этапа;

$Ordered_i$ – упорядоченный список задач для успешной реализации i-го спринта;

$Resplan_i$ – результат процесса планирования для задач, которые относятся к i-му спринту;

$Formteam_i$ – формирование командных задач для i-го спринта;

$Results_i$ – оценка итогов работы после окончания выполнения i-го спринта;

$Reaction_i$ – реакция заинтересованных сторон на полученные результирующие характеристики выполненного i-го спринта.

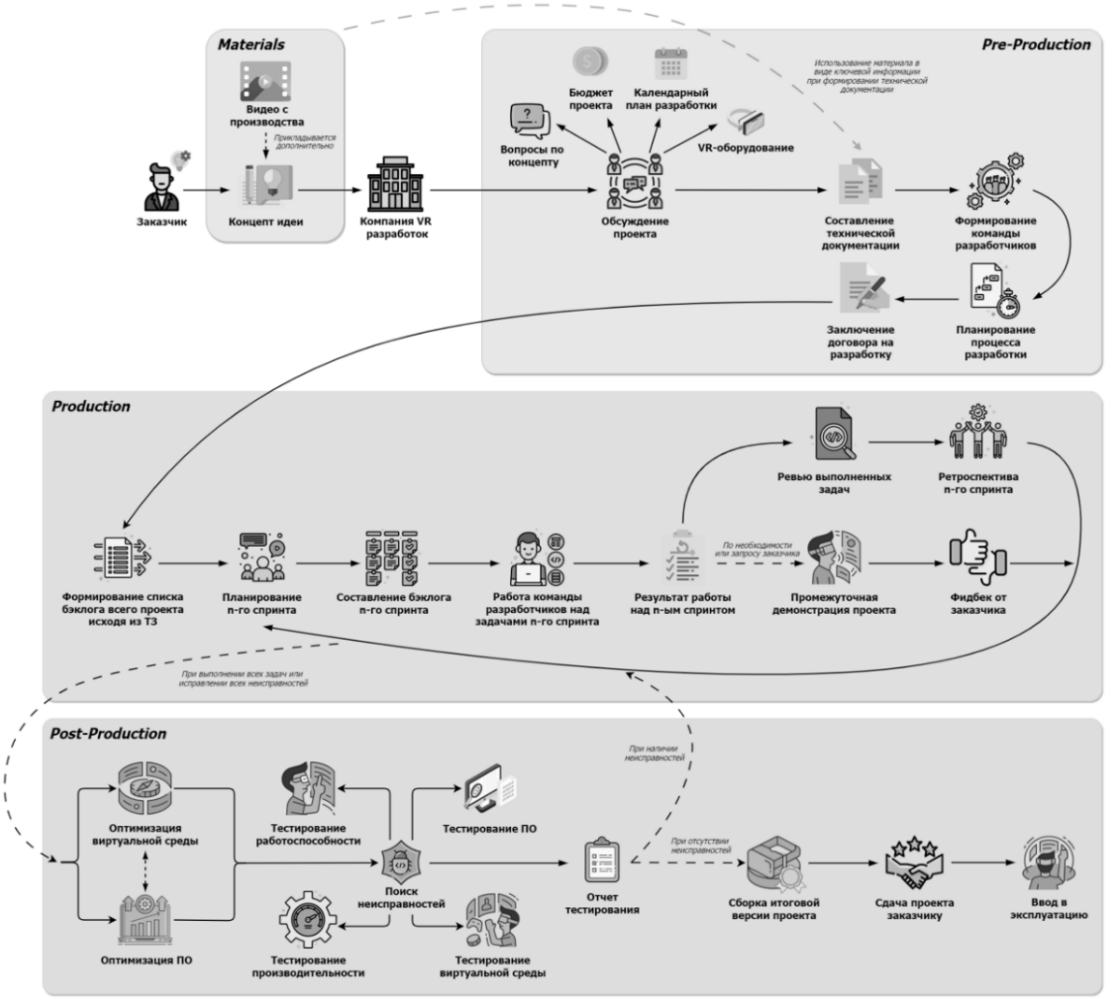


Рисунок 4 – Схема построения иммерсивных тренажеров для предметной области подготовки студентов инженерных специальностей

Завершающая фаза проекта сосредоточена на комплексном тестировании созданного решения. Осуществляется оценка производительности системы, проверка функциональной полноты, выявление и устранение программных дефектов. По результатам тестирования выполняется финальная оптимизация продукта, далее выполняется его передача заказчику совместно с требуемой сопроводительной документацией:

$$Testproject_{delivery} = v\left(Proctest_{perform}, Testfunc_{simul}, Ident_{corr}, Project_{delivery}, Project_{delivery}\right), \quad (4)$$

где $Proctest_{perform}$ – осуществление процесса тестирования достигнутых значений характеристик производительности проекта;

$Testfunc_{simul}$ – выполнение тестирования функциональных характеристик иммерсивного тренажера;

$Ident_{corr}$ – выявление ошибок и их последующее исправление;

$Finoptimiz_{simul}$ – заключительная оптимизация программного средства иммерсивного тренажера;

$Protect_{transf}$ – защита и передача проекта его заказчику.

Предложенная на основе бизнес-процессов модель ориентирована на комплексное построение иммерсивных тренажеров, причем рассмотренные этапы метода используют полученные на предыдущем результаты. Разработанный метод обеспечивает снижение рисков проектирования тренажеров, сокращение затрат, соответствие функциональных показателей программного продукта заданным условиям.

В целом процесс проектирования можно представить выражением:

$$Project\ Build = Pr_{pr} \rightarrow Project_{phase} \rightarrow PTestproject_{delivery}, \quad (5)$$

или в развернутом виде:

$$Project\ Build = f\left(Develop_{\text{concept}}, Plan_{budget}, Develop_{project}, Group_{specialists}, Implementation_{project}\right) \rightarrow \sum_{i=1}^{Ns} Iteration_i\ (Ordered_i, Resplan_i, Formteam_i, Results_i, Feedback_{cust}) \rightarrow v\left(Proctest_{perform}, Testfunc_{simul}, Ident_{corr}, Project_{delivery}, Project_{delivery}\right). \quad (6)$$

ЗАКЛЮЧЕНИЕ

Проведенное исследование позволило разработать комплексный метод проектирования симуляторов виртуальной реальности для инженерного образования, основанный на онтологическом моделировании предметной области. Ключевым результатом работы является создание формализованной онтологической модели, обеспечивающей структурное описание взаимосвязей между компонентами образовательного процесса и позволяющей унифицировать подход к разработке VR-тренажеров. Предложенная трехуровневая модель жизненного цикла разработки демонстрирует системный подход к созданию образовательных симуляторов. Особую значимость имеет формализация процесса сбора и анализа производственных видеоматериалов, обеспечивающая аутентичность создаваемых виртуальных сред.

Практическая апробация метода в рамках разработки симуляторов для подготовки студентов машиностроительных специальностей подтвердила его эффективность и воспроизводимость. Внедрение предложенных решений позволяет существенно повысить качество образовательных VR-продуктов на основе обеспечения их соответствия требованиям профессиональных стандартов и образовательных программ. Перспективы дальнейших исследований заключаются в расширении онтологической модели для включения смежных инженерных дисциплин, а также в разработке автоматизированных инструментов поддержки предложенного метода проектирования.

Список источников

1. Архипов, А. Е. Проектирование системы визуализации тренажерного комплекса на основе компетентностного подхода / А. Е. Архипов, А. И. Попов, А. Д. Обухов // Вопросы журналистики, педагогики, языкознания. – 2020. – Т. 39, № 3. – С. 378–390. – EDN KWOLEA.

2. Асланов, Р. Э. Виртуальные тренажеры в подсистеме АСУП для подготовки операторов металлорежущих станков: специальность 2.3.3. – Автоматизация и управление технологическими процессами и производствами : дис. … канд. техн. наук / Роман Эдвинович Асланов. – 2024. – 155 с.

3. Асланов, Р. Э. Использование симуляторов виртуальной реальности для подготовки операторов токарной и фрезерной обработок в подсистемах подготовки персонала АСУП / Р. Э. Асланов, А. А. Большаков. – Санкт-Петербург : Изд-во Политехн. ун-та, 2024. – 264 с. – ISBN 978-5-7422-8374-4.

4. Асланов, Р. Э. Модели и методы разработки подсистемы подготовки специалистов автоматизированной системы управления производством с использованием симуляторов виртуальной реальности / Р. Э. Асланов, А. А. Большаков // Известия Санкт-Петербургского государственного технологического института (технического университета). – 2023. – № 65 (91). – С. 81–89. – EDN EETIKD.

5. Асланов, Р. Э. Построение онтологической модели в виде семантической сети подсистемы подготовки персонала АСУП / Р. Э. Асланов, А. А. Большаков // Математические методы в технологиях и технике. – 2023. – № 10. – С. 17–23. – EDN STNQRN.

6. Бахтадзе, Н. Н. Предтренажерное обучение операторов сложных технических систем: инженерно-психологическое обоснование, методики, модели, цифровые решения / Н. Н. Бахтадзе, В. М. Дозорцев, А. А. Обознов, Е. Д. Чернецкая // Институт психологии Российской академии наук. Организационная психология и психология труда. – 2025. – Т. 10, № 1. – С. 173–201. – EDN HPSLDQ.

7. Бова, В. В. Компьютерная онтология: задачи и методология построения / В. В. Бова, Д. В. Лещанов, Д. Ю. Кравченко, А. А. Новиков // Информатика, вычислительная техника и инженерное образование. – 2014. – № 4 (19). – С. 44–55. – EDN UBNFFF.

8. Вешнева, И. В. Технологии виртуальной и дополненной реальности в образовании: проблемы, перспективы, реализация / И. В. Вешнева, А. А. Большаков, Р. Э. Асланов, Е. И. Малько // Математические методы в технологиях и технике. – 2025. – № 3. – С. 124–134. – EDN HKUPXH.

9. Вяльцев, А. В. Роль симуляционных технологий в образовательном процессе по обеспечению технологической безопасности на производстве / А. В. Вяльцев // Управление образованием: теория и практика. – 2024. – № 7–1. – С. 178–185. – EDN FKZHHV.

10. Ибрагимов, А. Р. Методологические аспекты разработки виртуальных тренажеров для обучения техническим навыкам: на примере проверки и обслуживания датчиков переменного перепада давления / А. Р. Ибрагимов, М. Г. Баширов, О. Г. Волкова // Наука и бизнес: пути развития. – 2025. – № 4 (166). – С. 40–46. – EDN ZYIIQY.

11. Иванов, М. Н. Виртуальные тренажеры как неотъемлемая составляющая электронного обучения / М. Н. Иванов // Мягкие измерения и вычисления. – 2020. – Т. 36, № 11. – С. 67–76. – DOI 10.36871/2618-9976.2020.11.006. – EDN DIVVJQ.

12. Кривов, М. В. Концепция синтеза компьютерных тренажерных комплексов для подготовки операторов / М. В. Кривов, А. Г. Колмогоров, В. Ю. Кобозев, Н. С. Благодарный // Вестник Ангарского государственного технического университета. – 2020. – № 14. – С. 104–108. – EDN FFRMMW.

13. Кузенко, С. Е. Разработка виртуального тренажера для подготовки электротехнического и электротехнологического персонала / С. Е. Кузенко, Л. Р. Вахитова, В. В. Галкина // Наука и бизнес: пути развития. – 2025. – № 4 (166). – С. 81–84. – EDN ROMEIT.

14. Ланских, Ю. В. Анализ и моделирование современных технологий формирования учебно-методических материалов / Ю. В. Ланских Ю.В. // ИТ Арктика. – 2021. – № 3. – С. 71–88. – EDN YBCCFT.

15. Печорин, А. С. Проектирование и разработка систем виртуальной реальности для образования и производства // А. С. Печорин, А. Г. Михайлова, М. А. Суворов, Н. О. Сулевич, В. В. Авсиевич // Мехатроника, автоматизация и управление на транспорте : материалы VI Всероссийской научно-практической конференции. – Самара, 2024. – С. 170–174. – EDN ULZRZR.

16. Случанинов, Н. Н. Виртуальные тренажеры как элемент информационных систем инженерного вуза / Н. Н. Случанинов, С. Д. Чижиумов // Специальная техника и технологии транспорта. – 2022. – № 14. – С. 348–353. – EDN PPSBWM.

17. Сметюх Н. П. Многофункциональные виртуальные тренажеры для подготовки экипажей судов рыбопромыслового флота: специальность 05.13.06 – Автоматизация и управление технологическими процессами и производствами (по отраслям) : дис. … канд. техн. наук / Надежда Павловна Сметюх. – 2017. – 185 с. – EDN CRMUSG.

18. Халикова, А. И. Методология разработки промышленных обучающих систем на примере компьютерного тренажера для операторов сборочных станков шинных производств / А. И. Халикова, М. Л. Шустрова, Н. А. Староверова // Южно-Сибирский научный вестник. – 2022. – № 6 (46). – С. 13–22. – EDN CAAUAM.

19. Штырлов, Ю. В. Моделирование виртуального тренажера на базе автоматного подхода / Ю. В. Штырлов // Инженерный вестник Дона. – 2025. – № 4 (124). – С. 803–814. – EDN TULEKK.

20. Aslanov, R. Method for constructing virtual reality simulators for turning and milling for an engineering education system for building cyber-physical systems / R. Aslanov, A. Bolshakov // Studies in Systems, Decision and Control. Cyber-Physical Systems: Society 5.0. Cyber-Solutions for Human-Centric Technologies. Springer Nature Switzerland AG 2023. – 2023. – Vol. 477. – P. 91–106. – https://doi.org/10.1007/978-3-031-35875-3_8.

21. Belyaev, S. A. Development of a multi-user software simulator based on web-technologies / S. A. Belyaev, V. S. Ivanov // Software Journal: Theory and Applications. – 2020. – № 2. – P. 1. – EDN DRODWI.

22. Doolani, S. An Immersive Virtual Storytelling System for Vocational Training / S. Doolani, L. Owens, C. Wessels, F. Makedon // Appl. Sci. – 2020. – Vol. 10. – P. 8143.

23. Guinn, I. V. Application of the new technologies: Augmented Reality and Virtual Reality in Education / I. V. Guinn // Cross-Cultural Studies: Education and Science. – 2022. – Vol. 7, iss. 2. – P. 126–132.

24. Lerner, D. An Immersive Multi-User Virtual Reality for Emergency Simulation Training: Usability Study / D. Lerner, S. Mohr, J. Schild, M. Göring, T. Luiz // JMIR Serious Games. – 2020. – Vol. 8, iss. 3.

References

1. Arkhipov, A. E., Popov, A. I., Obukhov, A. D. Design of a visualization system for a training complex based on a competency-based approach. *Issues of Journalism, Pedagogy, Linguistics*, 2020, vol. 39, no. 3, pp. 378–390 (In Russ.).

2. Aslanov, R. E. *Virtual simulators in the automated control system subsystem for training machine tool operators: specialty 2.3.3. – Automation and control of technological processes and production : dissertation for the degree of candidate of technical sciences*, 2024. 155 p. (In Russ.).

3. Aslanov, R. E., Bolshakov, A. A. *Using virtual reality simulators for training turning and milling operators in automated control system personnel training subsystems.* St. Petersburg, Publishing House of Polytechnical University, 2024. 264 p. ISBN 978-5-7422-8374-4 (In Russ.).

4. Aslanov, R. E., Bolshakov, A. A. Models and methods for developing a subsystem for training specialists of an automated production control system using virtual reality (In Russ.).

5. Aslanov, R. E., Bolshakov, A. A. Construction of an ontological model in the form of a semantic network of the personnel training subsystem of the automated control system. *Mathematical Methods in Technology and Engineering*, 2023, no. 10, pp. 17–23 (In Russ.).

6. Bakhtadze, N. N., Dozortsev, V. M., Oboznov, A. A., Chernetskaya, E. D. Pre-training training for operators of complex technical systems: engineering and psychological rationale, methods, models, and digital solutions. *Institute of Psychology of the Russian Academy of Sciences. Organizational and Occupational Psychology*, 2025, vol. 10, no. 1, pp. 173–201 (In Russ.).

7. Bova, V. V., Leshchanov, D. V., Kravchenko, D. Yu., Novikov, A. A. Computer ontology: tasks and methodology of construction. *Computer Science, Computing and Engineering Education*, 2014, no. 4 (19), pp. 44–55 (In Russ.).

8. Veshneva, I. V., Bolshakov, A. A., Aslanov, R. E., Malko, E. I. Virtual and augmented reality technologies in education: problems, prospects, implementation. *Mathematical Methods in Technology and Engineering*, 2025, no. 3, pp. 124–134 (In Russ.).

9. Vyaltsev, A. V. The role of simulation technologies in the educational process to ensure technological safety in production. *Education Management: Theory and Practice*, 2024, no. 7–1, pp. 178–185 (In Russ.).

10. Ibragimov, A. R., Bashirov, M. G., Volkova, O. G. Methodological aspects of developing virtual simulators for teaching technical skills: the example of testing and servicing differential pressure sensors. *Science and Business: Paths of Development*, 2025, no. 4 (166), pp. 40–46 (In Russ.).

11. Ivanov, M. N. Virtual simulators as an integral component of e-learning. *Soft Measurements and Calculations*, 2020, vol. 36, no. 11, pp. 67–76. DOI 10.36871/2618-9976.2020.11.006 (In Russ.).

12. Krivov, M. V., Kolmogor, A. G., Kobozev, V. Yu., Blagodarny, N. S. The concept of synthesis of computer training systems for operator training. *Bulletin of Angarsk State Technical University*, 2020, no. 14, pp. 104–108 (In Russ.).

13. Kuzenko, S. E., Vakhitova, L. R., Galkina, V. V. Development of a virtual simulator for training electrical engineering and electrotechnical personnel. *Science and Business: Paths of Development*, 2025, no. 4 (166), pp. 81–84 (In Russ.).

14. Lanskikh, Yu. V. Analysis and modeling of modern technologies for the formation of educational and methodological materials. *IT Arctic*, 2021, no. 3, pp. 71–88 (In Russ.).

15. Pechorin, A. S., Mikhailova, A. G., Suvorov, M. A., Sulevich, N. O., Avsievich, V.V. Design and development of virtual reality systems for education and production. *Mechatronics, Automation, and Control in Transport. Proceedings of the VI All-Russian Scientific and Practical Conference.* Samara, 2024, pp. 170–174 (In Russ.).

16. Sluchaninov, N. N., Chizhiumov, S. D. Virtual simulators as an element of information systems of an engineering university. *Specialized Equipment and Transport Technologies*, 2022, no. 14, pp. 348–353 (In Russ.).

17. Smetyukh, N. P. *Multifunctional virtual simulators for training crews of fishing vessels : specialty 05.13.06 – Automation and Control of Technological Processes and Production (by industry) : dissertation for the degree of Candidate of Technical Sciences*, 2017. 185 p. (In Russ.).

18. Khalikova, A. I., Shustrova, M. L., Staroverova, N. A. Methodology for developing industrial training systems using the example of a computer simulator for operators of tire assembly machines. *South Siberian Scientific Bulletin*, 2022, no. 6 (46), pp. 13–22 (In Russ.).

19. Shtyrlov, Yu. V. Modeling a virtual simulator based on the automata approach. *Engineering Bulletin of the Don*, 2025, no. 4 (124), pp. 803–814.

20. Aslanov, R., Bolshakov, A. Method for constructing virtual reality simulators for turning and milling for an engineering education system for building cyber-physical systems. Studies in Systems, Decision and Control. *Cyber-Physical Systems: Society 5.0. Cyber-Solutions for Human-Centric Technologies. Springer Nature Switzerland AG 2023*, 2023, vol. 477, pp. 91–106. https://doi.org/10.1007/978-3-031-35875-3_8.

21. Belyaev, S. A., Ivanov, V. S. Development of a multi-user software simulator based on web-technologies. *Software Journal: Theory and Applications*, 2020, no. 2, p. 1.

22. Doolani, S., Owens, L., Wessels, C., Makedon, F. An Immersive Virtual Storytelling System for Vocational Training. *Appl. Sci.*, 2020, vol. 10, p. 8143.

23. Guinn, I. V. Application of the new technologies: Augmented Reality and Virtual Reality in Education. *Cross-Cultural Studies: Education and Science*, 2022, vol. 7, iss. 2, pp. 126–132.

24. Lerner, D., Mohr, S., Schild, J., Göring, M., Luiz, T. An Immersive Multi-User Virtual Reality for Emergency Simulation Training: Usability Study. *JMIR Serious Games*, 2020, vol. 8, iss. 3.

Статья поступила в редакцию 21.10.2025; одобрена после рецензирования 25.11.2025; принята к публикации 26.11.2025.

The article was submitted 21.10.2025; approved after reviewing 25.11.2025; accepted for publication 26.11.2025.

ПРИБОРОСТРОЕНИЕ, МЕТРОЛОГИЯ И ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ ПРИБОРЫ И СИСТЕМЫ

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ И УПРАВЛЯЮЩИЕ СИСТЕМЫ

УДК 681.518.3

ПОСТРОЕНИЕ ИЗМЕРИТЕЛЬНЫХ КАНАЛОВ В ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ МОНИТОРИНГА ФИЗИЧЕСКИХ ПРОЦЕССОВ

Григорьян Леонтий Рустемович, Кубанский государственный университет, 350040, Российская Федерация, г. Краснодар, ул. Ставропольская, 149,
кандидат физико-математических наук, доцент, ORCID: 0009-0002-4744-0356, e-mail: leonmezon@mail.ru
Богатов Николай Маркович, Кубанский государственный университет, 350040, Российская Федерация, г. Краснодар, ул. Ставропольская, 149,
доктор физико-математических наук, профессор, ORCID: 0000-0002-9301-4545, e-mail: bogatov@phys.kubsu.ru
Коваленко Максим Сергеевич, Кубанский государственный университет, 350040, Российская Федерация, г. Краснодар, ул. Ставропольская, 149,
кандидат физико-математических наук, доцент, ORCID: 0009-0006-5402-7366, e-mail: m.s.kovalenko@ya.ru
Сахно Мария Александровна, ООО «Северо-Западный региональный центр аэронавигационной информации», 190121, Российская Федерация, г. Санкт-Петербург, наб. реки Пряжки, 18–20, лит. А, пом. 2Н,
ведущий проектный менеджер, ORCID: 0009-0004-7462-9773, e-mail: maria_210492@mail.ru

В работе рассматривается задача построения системы сбора данных физических процессов с учетом нахождения оптимальной схемотехнической и программной составляющей построения системы мониторинга, а также обеспечения непрерывного процесса измерения сигналов с использованием оптимальных параллельных алгоритмов обработки аналого-цифровой информации. Для построения аналоговых и аналого-цифровых измерительных трактов сигналов многоканальной системы предложены оптимальные схемотехнические решения, основываясь на требовании полноты физической информации. Рассмотрена архитектура управляющей системы мониторинга, предложены оригинальные подходы к ее программной составляющей как по компонентам, так и по взаимосвязи.
Ключевые слова: датчики, физические процессы, системы сбора данных, измерение, распараллеливание

CONSTRUCTION OF MEASURING CHANNELS IN INFORMATION AND MEASURING SYSTEMS FOR MONITORING PHYSICAL PROCESSES

Grigoryan Leonty R., Kuban State University, 149 Stavropolskaya St., Krasnodar, 350040, Russian Federation,
Cand. Sci. (Physics and Mathematics), Associate Professor, ORCID: 0009-0002-4744-0356, e-mail: leonmezon@mail.ru
Bogatov Nikolai M., Kuban State University, 149 Stavropolskaya St., Krasnodar, 350040, Russian Federation,
Doct. Sci. (Physics and Mathematics), Professor, ORCID: 0000-0002-9301-4545, e-mail: bogatov @phys.kubsu.ru
Kovalenko Maksim S., Kuban State University, 149 Stavropolskaya St., Krasnodar, 350040, Russian Federation,
Cand. Sci. (Physics and Mathematics), Associate Professor, ORCID: 0009-0006-5402-7366, e-mail: m.s.kovalenko@ya.ru
Sakhno Maria A., “North-West Regional Aeronautical Information Center” LLC, 18–20, lit. A, room 2N Pryazhka River Embankment, St. Petersburg, 190121, Russian Federation,
leading project manager, ORCID: 0009-0004-7462-9773, e-mail: maria_210492@mail.ru

This paper examines the problem of constructing a system for collecting data on physical processes, considering the optimal circuitry and software components for constructing a monitoring system, as well as ensuring continuous signal measurement using optimal parallel algorithms for processing analog-to-digital information. Optimal circuitry solutions are proposed for constructing analog and analog-to-digital signal measurement paths for a multichannel system, based on the requirement for completeness of physical information. The architecture of the control monitoring system is examined, and original approaches to its software component, both in terms of components and interconnections, are proposed.
Keywords: sensors, physical processes, data acquisition systems, measurement, parallelization

ВВЕДЕНИЕ

Научно-технические решения, применяемые в познавательной деятельности человека, с каждым годом охватывают все больший круг разнообразных применений. Они особенно характерны для тех объектов, которые распределены на некоторой территории или для которых постоянное присутствие человека по тем или иным причинам нежелательно или экономически нецелесообразно. В этих случаях дистанционное автоматическое диспетчерское управление и контроль являются необходимыми.

В процессе познавательной деятельности возникают практические задачи, требующие количественной оценки информации, анализ которой позволяет в дальнейшем принимать определенные решения, при этом процесс измерения является основополагающим способом получения количественной информации [1]. Следует отметить, что принимаемые решения на основе обработки информации в любых системах будут истинны только в том случае, если процесс получения первичных данных будет строго объективен и достоверен. Поэтому с увеличением разнообразия практического применения научно-технических решений растут и требования к построению измерительных систем как в плане повышения их точности, так и достоверности процесса измерения [2]. В то же время существует достаточно широкий набор средств, которые позволяют создавать сложные системы измерения, контроля и мониторинга физических процессов, но их конфигурация пластична в зависимости от задачи.

Цель работы: проектирование измерительных систем с параллельной конфигурацией контроля физических параметров.

Определение цели позволяет сформулировать основные задачи, требующие своего освещения:

– анализ основных функциональных блоков многоканального измерительного модуля;
– анализ возможности совмещения способов аналоговой и цифровой фильтрации сигналов;
– определение основных способов внутренней синхронизации измерительного модуля;
– определение основных критериев, обеспечивающих быстродействие цифрового измерительного блока.

АРХИТЕКТУРА ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНОЙ СИСТЕМЫ

В структуре параллельных измерительных систем чаще всего используют многоканальные автоматизированные измерительные системы (кроме ограниченного круга задач, требующих использования единственного высокоскоростного измерительного канала и обусловливающих применение мощных и быстродействующих вычислительных ресурсов) [3]. Основные достоинства многоканальных измерительных систем связаны с возможностью измерения разнородных физических величин, достижения максимального быстродействия и высокой схемной надежности [4].

Практическая реализация многоканальной автоматизированной системы строится на основе N каналов измерения и в общем случае представляет собой наиболее сложный узел, так как определяет метрологические характеристики всей системы [5]. Типичная схема реализации измерительного канала приведена на рисунке 1.

Рисунок 1 – Функциональная схема реализации канала измерения

Данные с датчика физических величин $f(x)$ поступают на масштабный преобразователь (например, усилитель), назначение которого связано с преобразованием уровня аналогового сигнала до приемлемого значения. Далее аналоговый сигнал поступает на устройство фильтрации, которое в общем случае может состоять из разных по назначению фильтров, построенных как по аналоговой схеме, так и по цифровой. После фильтрации сигнал поступает на аналого-цифровой преобразователь, выполняющий наиболее ответственную функцию дискретизации аналогового сигнала. И далее цифровой код $N(x)$, пропорциональный физической величине x, передается на управляющее устройство. Устройство управления канала измерения выполняет функцию контроля и управления входящими в канал устройствами. Такое построение позволяет достаточно просто реализовать канал измерения, однако не обеспечивает решения всего спектра возможных измерительных задач [6].

Действительно, использование раздельного управления в каждом канале измерения, с одной стороны, требует усложнения общего управления всеми каналами измерения (требует настройки

каждого канала измерения по отдельности) [7], с другой стороны, не позволяет точно синхронизировать измерения по каналам, и, в-третьих, отсутствие промежуточных вычислительных процессоров не позволяет всей автоматизированной системе в реальном времени реагировать на изменение входной информации [8].

Одним из возможных решений указанной выше задачи является построение многоканальных измерительных систем, основанных на программно-аппаратных алгоритмах обработки информации с возможностью точной синхронизации процесса дискретизации по времени. Рассмотрим функциональную схему многоканальной измерительной системы (рис. 2).

Измеряемые физические параметры $f(x_1), f(x_2)$ и $f(x_i)$ с датчиков поступают на масштабные преобразователи сигналов. Параметры масштабирования сигналов задаются по линии управления, причем для каждого канала индивидуальные с учетом цифровой идентификации канала. Такая организация измерений позволяет настраивать каждый канал по отдельности, согласно уровню исследуемого им сигнала, что позволяет организовать исследование сигналов в широком динамическом диапазоне.

Далее сигналы поступают на устройства фильтрации, которые аналогично подключены к линии управления с цифровым заданием параметров [9]. В качестве управляемых фильтров используются перенастраиваемые аналоговые фильтры как с управлением на аналоговых ключах, так и полностью аналоговые с цифровым управлением. Такое решение позволяет не только выполнять стандартные операции фильтрации сигнала, но и, применяя специализированные алгоритмы обработки информации, оперативно реагировать на изменяющиеся помехи и одновременно адаптировать параметры фильтрации, что существенно расширяет возможности всей многоканальной автоматизированной системы и увеличивает общую помехоустойчивость процесса измерения сигналов.

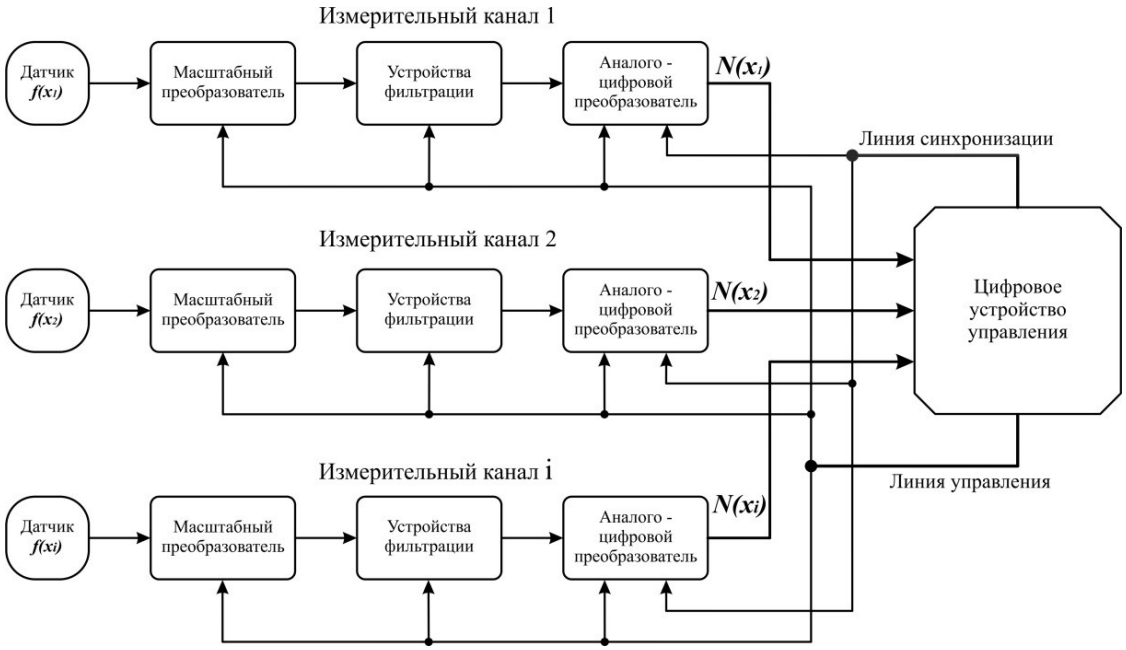


Рисунок 2 – Функциональная схема многоканальной измерительной системы

Отфильтрованные данные поступают далее на аналого-цифровые преобразователи, управление которыми осуществляется единым цифровым устройством по двум линиям: управления и синхронизации, что позволяет решить задачу синхронной дискретизации сигналов по времени. При этом точность синхронизации будет определяться в первую очередь точностью устройства синхронизации, в качестве которого могут использоваться как классические кварцевые источники высокоточных синхроимпульсов, так и генераторы синхроимпульсов, реализованные с использованием спутниковой глобальной системы позиционирования.

Применение единого цифрового устройства управления в целом повышает помехоустойчивость системы и одновременно обеспечивает синхронизацию по времени поступающих данных. В качестве цифровых устройств могут использоваться микроконтроллеры, микропроцессоры и в некоторых случаях цифровые сигнальные процессоры, что в целом обеспечивает высокую гибкость настройки системы и достоверность обработки измерительных сигналов.

Следует также отметить, что использование цифрового устройства управления не ограничивается рассмотренными выше задачами и его применение можно значительно расширить в аспектах

обработки и интерпретации исследуемых данных [10]. Например, цифровая программная фильтрация в цифровом устройстве позволяет дополнительно провести обработку исследуемой измерительной информации и целенаправленно изменить соотношение между различными компонентами спектра сигнала. Целью такой обработки может быть подавление помех либо выделение составляющих сигнала, которые соответствуют определенным свойствам исследуемого процесса [11].

Реализация программного блока цифровой обработки сигналов определяется как выбором языка программирования с учетом необходимой скорости обработки, так и заданными аппаратными возможностями цифрового устройства управления. Это определяется тем, что при измерении уровня первой гармоники в анализе спектра сигнала определяющим является качество цифровой фильтрации гармоник сигнала от первой до заданной (например, сороковой) [12]. Таким образом, выбор структуры цифрового устройства управления, в плане цифровой обработки сигналов, определяется требованиями непосредственно алгоритма цифровой обработки сигналов.

Добавление в архитектуру цифрового устройства возможности управления дополнительными устройствами преобразует систему из измерительной в управляющую. Это накладывает определенные требования на выбор архитектуры цифрового устройства управления как по быстродействию, так и по вычислительным емкостным ресурсам, а также по возможности управления дополнительными подключенными исполнительными устройствами. Чаще всего цифровое устройство управления напрямую невозможно подключить к исполнительным устройствам без дополнительных буферных блоков из-за энергетических (токовых) ограничений. Кроме того, использование цифровых линий потребует применения соответствующих дешифрирующих устройств.

Измеренная и обработанная информация в целом может являться входными данными для другой системы, более высокого уровня, что потребует организации передачи первично обработанных данных. В этом случае измерительная система может быть одной из составляющих большой информационно-измерительной системы, что потребует наличия соответствующих каналов связи и дополнительных вычислительных ресурсов при реализации алгоритмов трансформации информации. Данные возможности определяются выбором аппаратных и программных ресурсов цифрового устройства управления и требований по сопряжению с системой верхнего уровня.

Задача визуального отображения измеренной информации для оперативного контроля процесса измерения физических величин требует использования показывающих устройств с собственным аппаратно-программным блоком преобразования информации, а также с использованием возможностей цифрового устройства управления.

Приведенные характеристики цифрового устройства управления позволяют сформировать обобщенную архитектуру измерительной системы, представленную на рисунке 3.

Рисунок 3 – Архитектура программного обеспечения измерительной системы

Архитектура программного обеспечения измерительной системы состоит из следующих модулей: модуль общей синхронизации, модуль управления аналого-цифровыми преобразователями, модуль связи, модуль цифровой обработки сигналов, общий модуль управления, модуль управления прерываниями, модуль управления таймерами, модуль вывода информации.

Рассмотрим далее структуру и назначение обозначенных модулей, общее управление которыми осуществляется управляющими сигналами от базового программного модуля управления.

Модуль управления аналого-цифровыми преобразователями задает параметры и принимает оцифрованные сигналы, осуществляя преобразование данных в заданный диапазон. Модуль цифровой обработки сигналов осуществляет фильтрацию оцифрованных сигналов. Поступающий сигнал не всегда удовлетворяет запрашиваемым требованиям, поэтому из него необходимо выделить ту часть сигнала, которая непосредственна необходима. Модуль связи с показывающими устройствами обеспечивает как отображение основных параметров работы измерительной системы, так и вывод на показывающие устройства. Модуль связи с системой верхнего уровня предназначен для приема и передачи данных по каналам связи от системы верхнего уровня, модуль также осуществляет шифрование и расшифровку передаваемых и принимаемых данных. Модуль общей синхронизации синхронизирует такты работы аналого-цифровых преобразователей, фронты внешних прерываний и прерываний от таймеров-счетчиков, а также начало передачи данных через модуль связи. Тактирование модуля синхронизации осуществляется как от кварцевого генератора, так и от сигналов глобальной системы позиционирования, а в некоторых случаях и от системы верхнего уровня. Модуль управления прерываниями осуществляет задание параметров и выполнение соответствующих процедур при возникновении события в регистрах внешних прерываний. Модуль управления таймерами выполняет функцию задания параметров таймеров-счетчиков, работающих на точность дискретизации аналого-цифровых преобразователей, и тактируется от модуля общей синхронизации. Модуль управления фильтрами и усилителями предназначен для согласования амплитудных параметров датчиков физических величин с динамическим диапазоном аналого-цифровых устройств. Модуль управления исполнительными устройствами является необязательным и дополнительным. Модуль может осуществлять воздействие на объект исследования при организации обратной связи или являться базовым управляющим механизмом подчиненных устройств [13].

В целом работоспособность всей системы задается операционной системой реального времени (Embedded Linux или Windows IoT) [14]. Это позволяет добиться высокой частоты измерений и эффективной обработки быстро меняющихся сигналов. Гибкость в функционале достигается модульной архитектурой, реализация которой возможна средствами языка C++ без ущерба производительности системы. При этом обеспечивается низкая связанность между модулями системы и быстрая замена или модификация функционала отдельно взятого модуля без риска нарушить функционирование всей системы в целом. Для каждого типа модулей может быть несколько программных реализаций под конкретные задачи. Каждая реализация может использовать алгоритмы и библиотеки, наиболее эффективно решающие необходимые задачи. Например, Eigen – для фильтрации сигнала и выделения значимых для мониторинга составляющих, FFTW – для спектрального анализа, Dlib и mlpack – для обнаружения аномалий и интеллектуальной проверки корректности данных в реальном времени [15]. Кроме того, допускается расширение стандартного функционала модулей дополнительными возможностями. Например, модуль связи помимо передачи данных измеренных сигналов может выполнять автоматическое оповещение при отказе элементов измерительной системы или периодическую проверку работоспособности датчиков, повышая отказоустойчивость системы [16].

В качестве практического применения измерительной системы мониторинга физических процессов целесообразно привести работы авторов, в которых вопросы построения измерительной системы получили свое воплощение и специальное развитие:

- система сбора данных и управления пчелиным ульем [17];
- построение системы сбора данных геофизических сигналов [18];
- построение системы маркерной топографии [19].

ЗАКЛЮЧЕНИЕ

В работе рассмотрены решения построения измерительных каналов в информационно-измерительной системе мониторинга физических процессов.

Показано, что для повышения достоверности и эффективности обработки измерительных сигналов целесообразно в нескольких каналах измерения автоматизированных систем в качестве управляющего устройства применять микроконтроллеры совместно с устройствами синхронизации и оцифровки аналоговых данных. Это существенно расширит возможности измерительной системы, значительно повысит ее помехоустойчивость и обеспечит синхронизацию поступающих данных.

Требования к программному обеспечению метрологических систем изложены в [20, 21]. Так, гарантированное прохождение информации по измерительным каналам возможно с учетом разработки алгоритмов подготовки данных. При этом разработка алгоритмов шифрования позволяет увеличивать количество исследователей, что в целом гарантирует надежность управления всей системой [21].

Список источников

1. Левенец, А. В. Информационные процессы и системы. Обработка данных / А. В. Левенец. – Хабаровск : Тихоокеанский государственный университет, 2019. – 130 с.

2. Билашова, Г. И. Современные автоматизированные измерительные информационные системы / Г. И. Билашова // Наука через призму времени. – 2023. – № 2 (71). – С. 9–11.

3. Аманкаев, Ю. С. Оптимизация измерительных информационных систем / Ю. С. Аманкаев // Международная конференция по мягким вычислениям и измерениям. – 2017. – Т. 3. – С. 248–251.

4. Соколов, М. П. Автоматические измерительные устройства в экспериментальной физике / М. П. Соколов. – 2-е изд. – Москва : Атомиздат, 1978. – 351 с.

5. Юнусова, А. К. Прикладные функции информационных измерительных систем / А. К. Юнусова // Международная конференция по мягким вычислениям и измерениям. – 2017. – Т. 3. – 304–307 с.

6. Базыкин, С. Н. Концепция работы интеллектуальных датчиков в распределенных информационно-измерительных и управляющих системах / С. Н. Базыкин, В. А. Бардин, П. С. Чернов // Приборы. – 2023. – № 12 (282). – С. 1–7.

7. Полтавский, А. В. Особенности построения информационных и информационно-измерительных систем подготовки принятия управленческих решений / А. В. Полтавский // Двойные технологии. – 2024. – № 3 (108). – С. 60–63.

8. Волков, В. Л. Измерительные информационные системы / В. Л. Волков. – Арзамас : АПИ НГТУ, 2008. – 158 с.

9. Солонина, А. И. Алгоритмы и процессы цифровой обработки сигналов / А. И. Солонина, Д. А. Улахович, Л. А. Яковлев. – Санкт-Петербург : БХВ-Петербург, 2015. – 464 с.

10. Клочков, А. Е. Коммутаторы средств измерения – ключ к съему данных с цифровых средств измерения / А. Е. Клочков // Мир измерений. – 2025. – № 2. – С. 72–74.

11. Умняшкин, С. В. Основы теории цифровой обработки сигналов : учебное пособие / С. В. Умняшкин. – Изд. 7, испр. – Москва : Техносфера, 2024. – 552 с.

12. Нефедов, В. Г. Оценка эффективности алгоритмов цифровой обработки сигналов на ЭВМ / В. Г. Нефедов, Г. А. Петров, Е. В. Силяков, А. В. Семенов, А. Д. Филин // Флагман науки. – 2025. – № 1 (24). – С. 395–398.

13. Захаров, Н. А. Сетевые встраиваемые системы / Н. А. Захаров, В. И. Клепиков, Д. С. Подхватилин // Автоматизация в промышленности. – 2020. – № 3. – С. 58–61.

14. Устинов, А. В. Архитектура программного обеспечения распределенной систем обработки данных при автоматизированном анализе информации / А. В. Устинов, М. Ю. Охтилев // I-methods. – 2018. – Vol. 10. № 2. – С. 31–37.

15. Пруцков, А. В. Принципы разработки программных интерфейсов в индустриальных информационно-измерительных и управляющих системах / А. В. Пруцков // Контроль. Диагностика. – 2021. – Т. 24, № 10 (280). – С. 44–47.

16. Кривель, С. М. Методика и программное обеспечение моделирования функционирования систем с учетом характеристик их надежности на этапах эксплуатации и проектирования / С. М. Кривель, А. А. Лебедева, А. Б. Спасибко // Crede Experto: транспорт, общество, образование, язык. – 2023. – № 1. – С. 57–76.

17. Григорьян, Л. Р. Интеллектуальная система мониторинга пчелиных ульев / Л. Р. Григорьян, М. С. Коваленко, А. Л. Григорьян, Д. Ю. Парошин // Аграрный научный журнал. – 2019. – № 10. – С. 59–65.

18. Григорьян, Л. Р. Практика построения регистраторов геофизических сигналов / Л. Р. Григорьян, Н. М. Богатов, А. Л. Григорьян, В. В. Анисимова // Системы контроля окружающей среды. 2019. № 2 (36). С. 5–12.

19. Григорьян, Л. Р. Схемотехника электронных устройств маркерной топографической технологии Л. Р. Григорьян, Н. М. Богатов, Р. Л. Григорьян, М. А. Сахно // Приборы. – 2025.– № 3 (297). – С. 16–25.

20. Слаев, В. А. Аттестация программного обеспечения, используемого в метрологии : справочная книга / В. А. Слаев, А. Г. Чуновкина ; под ред. В. А. Слаева. – Санкт-Петербург : Профессионал, 2009. – 320 с.

21. Данилов, А. А. Направления совершенствования измерительных систем и их метрологического обеспечения / А. А. Данилов // Измерительная техника. – 2023. – № 8. – С. 24–29.

References

1. 1. Levenets, A. V. *Information processes and systems. Data processing.* Khabarovsk, Pacific State University, 2019. 130 p. (In Russ.).

2. Bilashova, G. I. Modern automated measuring information systems. *Science through the Prism of Time*, 2023, no. 2 (71), pp. 9–11 (In Russ.).

3. Amankaev, Yu. S. Optimization of measuring information systems. *International Conference on Soft Computing and Measurements*, 2017, vol. 3, pp. 248–251 (In Russ.).

4. Sokolov, M. P. *Automatic measuring devices in experimental physics*. 2nd ed. Moscow, Atomizdat Publ., 1978. 351 p. (In Russ.).

5. Yunusova, A. K. Applied functions of information measuring systems. *International Conference on Soft Computing and Measurement,* 2017, vol. 3, pp. 304–307 (In Russ.).

6. Bazykin, S. N., Bardin, V. A., Chernov, P. S. The concept of intelligent sensors in distributed information-measuring and control systems. *Devices*, 2023, no. 12 (282), pp. 1–7 (In Russ.).

7. Poltavsky, A. V. Features of the construction of information and information-measuring systems for preparing management decisions. *Dual Technologies*, 2024. no. 3 (108), pp. 60–63 (In Russ.).

8. Volkov, V. L. *Measuring information systems.* Arzamas, API NSTU, 2008. 158 p. (In Russ.).

9. Solonina, A. I., Ulakhovich, D. A., Yakovlev, L. A. *Algorithms and processes of digital signal processing*. Saint Petersburg, BHV-Petersburg Publ., 2015. 464 p. (In Russ.).

10. Klochkov, A. E. Switches of measuring instruments – the key to reading data from digital measuring instruments. *The world of Measurements*, 2025, no. 2, pp. 72–74 (In Russ.).

11. Umnyashkin, S. V. *Fundamentals of the theory of digital signal processing : tutorial.* 7 ed., rev. Moscow, Tekhnosfera Publ., 2024. 552 p. (In Russ.).

12. Nefedov, V. G., Petrov, G. A., Silyakov, E. V., Semenov, A. V., Filin, A. D. Evaluation of the effectiveness of digital signal processing algorithms on a computer. *Flagship of Science*, 2025, no. 1 (24), pp. 395–398 (In Russ.).

13. Zakharov, N. A., Klepikov, V. I., Podkhvatilin, D. S. Networked embedded systems. *Automation in Industry*, 2020, no. 3, pp. 58–61 (In Russ.).

14. Ustinov, A. V., Okhtilev, M. Yu. Software architecture of distributed data processing systems for automated information analysis. *I-methods*, 2018, vol. 10, no. 2, pp. 31–37 (In Russ.).

15. Prutskov, A. V. Principles of developing software interfaces in industrial information-measuring and control systems / *Control. Diagnostics*. 2021. vol. 24. no. 10 (280), pp. 44–47 (In Russ.).

16. Krivel, S. M., Lebedeva, A. A., Spasibko, A. B. Methodology and software for modeling the functioning of systems taking into account the characteristics of their reliability at the stages of operation and design. *Crede Experto: Transport, Society, Education, Language*. 2023, no. 1, pp. 57–76 (In Russ.).

17. Grigoryan, L. R., Kovalenko, M. S., Grigoryan, A. L., Paroshin, D. Y. Intelligent monitoring system for beehives. *Agricultural Scientific Journal*, 2019, no. 10, pp. 59–65 (In Russ.).

18. Grigoryan, L. R., Bogatov, N. M., Grigoryan, A. L., Anisimova, V. V. Practice of constructing geophysical signal recorders. *Environmental Monitoring Systems*, 2019, no. 2 (36), pp. 5–12 (In Russ.).

19. Grigoryan, L. R., Bogatov, N. M., Grigoryan, R. L., Sakhno, M. A. Circuitry of Electronic Devices of Marker Topographic Technology. *Devices*, 2025, no. 3 (297), pp. 16–25 (In Russ.).

20. Slaev, V. A. (ed.), Chunovkina, A. G. *Certification of software used in metrology: reference book*. St. Petersburg, Professional Publ., 2009. 320 p. (In Russ.).

21. Danilov, A. A. Directions for improving measuring systems and their metrological support. *Measuring Equipment*, 2023, no. 8, pp. 24–29 (In Russ.).

Статья поступила в редакцию 17.10.2025; одобрена после рецензирования 25.11.2025; принята к публикации 05.12.2025.

The article was submitted 17.10.2025; approved after reviewing 25.11.2025; accepted for publication 05.12.2025.

ПРАВИЛА ДЛЯ АВТОРОВ

1. В журнале публикуются материалы на английском и русском языках по тематике, соответствующей утвержденным для журнала отраслям наук, группам специальностей.

2. В список соавторов работ включаются только те лица, которые внесли творческий вклад в подготовку представленных материалов. Лицам, оказавшим только техническую помощь, можно выразить благодарность в конце статьи. Один человек может быть автором (соавтором) не более чем двух статей в одном номере журнала, причем единственным автором он может быть только в одной статье.

3. Объем публикаций для научных статей должен быть не менее 8 страниц, а количество источников в библиографическом списке (списке литературы) – не менее 10 позиций.

4. Содержание каждой статьи должно включать следующие элементы: УДК; название статьи; сведения об авторах, включая их место работы, должность, адрес электронной почты; аннотацию объемом от 100 до 250 слов, ключевые слова (от 9 до 13); графическую аннотацию, отражающую содержание статьи; название статьи, сведения об авторах, аннотацию и ключевые слова на английском языке (для англоязычных статей – на русском языке); введение – оно должно заканчиваться формулировкой цели работы в явной форме; собственно текст статьи – очень желательна его сегментация на разделы, имеющие содержательные заголовки; выводы или за- ключение (должны соответствовать формулировке цели статьи).

5. Для русскоязычных статей приводится два библиографических списка: на языке оригинала статьи; список с транслитерацией русскоязычных источников на латиницу и (дополнительно) приведением в квадратных скобках переводов названий статей и названий источников на английский язык.

В «русскоязычном» библиографическом списке (списке литературы) порядок следования источников – по алфавиту фамилий авторов (сначала русскоязычные источники, потом иноязычные). На все источники, включенные в библиографический список, должны быть даны ссылки в тексте статьи в квадратных скобках. При необходимости авторы могут указывать номера страниц в источниках, на которые даются ссылки. Приветствуются ссылки на иноязычные источники, а также на материалы, опубликованные ранее в журнале «Прикаспийский журнал: управление и высокие технологии». Однако в последнем случае количество таких ссылок не должно превышать 20 % от общего количества источников, включенных в библиографический список. Для источников, имеющих DOI, целесообразно его указывать. При ссылках на статьи, опубликованные в журнале «Прикаспийский журнал: управление и высокие технологии», целесообразно в конце библиографического описания источника в круглых скобках указывать гиперссылку, указывающую на место размещения статьи на страничке сайта Астраханского государственного университета.

Ссылки в библиографическом списке на материалы, размещенные в интернете, допускаются при соблюдении следующих условий: если у материала, на который дается ссылка, имеется автор и/или название, то они должны быть указаны для этого источника; должен быть приведен полный маршрут доступа к источнику в интернете; должна быть указана дата обращения (доступа) к источнику.

Ограничения по списку литературы: доля самоцитирований для любого из авторов статьи, а также по совокупности всех авторов статьи, не должна превышать 25 %; доля ссылок на статьи с участием одного автора, не являющегося автором (соавтором) статьи, не должна превышать 25 %.

6. Суммарная доля таблиц и иллюстраций в общем объеме представляемой статьи не должна превышать 40 %. Под иллюстрациями понимаются следующие объекты: диаграммы; графики; рисунки; эскизы; фотографии; карты и т.п.

7. Доля оригинального текста в статьях (оцениваемого через систему «Антиплагиат» на сайте www.antiplagiat.ru) должна быть не менее 80 %.

8. Указание на то, что работа финансируется по какому-либо гранту, в рамках Федеральной целевой программы, государственного заказа и пр. дается в виде постраничной сноски после заголовка (названия) работы.

9. В сведения об авторах работ помимо места работы и должности целесообразно включать ORCID автора и гиперссылку на страничку с его личными наукометрическими показателями на сайте www.elibrary.ru. По желанию можно привести также ссылки на странички с наукометрическими показателями на Scopus, в ResearchGate; на личную страничку, размещенную на сайте организации.

10. Основные технические требования к оформлению статей (материалов):

10.1. Текст должен быть расположен по ширине страницы формата А4 с учётом полей (все поля по 2,5 см), набран шрифтом Times New Roman, кегль 12, межстрочный интервал 1,0. В таблицах, подрисуночных надписях допускается уменьшенный шрифт – вплоть до 10 кегля. Альбомная ориентация страниц допускается только в порядке исключения для следующих случаев: широкоформатные таблицы с большим количеством колонок; иллюстрации большого размера, которые не умещаются на странице с книжной ориентацией.

Абзацные отступы одинаковы по всему тексту – 0,75 см. Кавычки («»), скобки ([ ], ( )), маркеры и другие знаки должны быть аналогичными на протяжении всего предоставляемого для публикации материала.

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2025
№ 4 (72)

Свидетельство о регистрации средства массовой информации
Федеральной службы по надзору в сфере массовых коммуникаций,
связи и охраны культурного наследия
ПИ № ФС77-31932 от 16 мая 2008 г.

Учредитель
Астраханский государственный университет имени В. Н. Татищева
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20а

Адрес редакции:
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20

Адрес издателя:
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20а

Редактирование,
компьютерная правка, верстка *Н. Н. Сахно*

Дата выхода в свет 20.02.2026 г.

Цена свободная
Уч.-изд. 12,0. Усл. печ. л. 16,8.
Заказ № 4727. Тираж 500 экз. (первый завод – 23 экз.)

Астраханский государственный университет имени В. Н. Татищева
414056, г. Астрахань, ул. Татищева, 20а
тел. (8512) 24-66-60 (доб. 3; издательско-полиграфический отдел)
Отпечатано в Астраханской цифровой типографии
414040, г. Астрахань, пл. К. Маркса, 33
тел./факс (8512) 54-00-11, 73-40-40,
E-mail: a-d-t@mail.ru