

Учредитель

ФГБОУ ВО «Новосибирский государственный технический университет»

Редакционный совет

Председатель редакционного совета

Ложников Павел Сергеевич, д-р техн. наук, доцент, НГТУ, г. Новосибирск

Заместители председателя

Белим Сергей Викторович, д-р физ.-мат. наук, проф., ОмГТУ, г. Омск

Вострецов Алексей Геннадьевич, д-р техн. наук, проф., засл. деятель науки РФ, НГТУ, г. Новосибирск

Котенко Игорь Витальевич, д-р техн. наук, проф., СПИИРАН, г. Санкт-Петербург

Члены редакционного совета

Авдеенко Татьяна Владимировна, д-р техн. наук, проф., НГТУ, г. Новосибирск

Алгулиев Расим Магомед, д-р техн. наук, проф., академик НАН Республики

Азербайджан, ИИТ МНО Азербайджанской Республики, г. Баку

Александрова Елена Борисовна, д-р техн. наук, доцент, СПбПУ, г. Санкт-Петербург

Аникин Игорь Вячеславович, д-р техн. наук, доцент, КНИТУ-КАИ, г. Казань

Арутюнян Мариам Евгеньевна, д-р физ.-мат. наук, проф., Институт проблем информатики и автоматизации НАН Республики Армения, г. Ереван

Бабенко Михаил Григорьевич, д-р физ.-мат. наук, доцент, СКФУ, г. Ставрополь

Баранкова Инна Ильинична, д-р техн. наук, доцент, МГТУ им. Г.И. Носова,

г. Магнитогорск

Беззатеев Сергей Валентинович, д-р техн. наук, доцент, СПбГУАП,

г. Санкт-Петербург

Васильев Владимир Иванович, д-р техн. наук, проф., УГАТУ, г. Уфа

Воевода Александр Александрович, д-р техн. наук, проф., НГТУ, г. Новосибирск

Вострецов Алексей Геннадьевич, д-р техн. наук, проф., засл. деятель науки РФ, НГТУ, г. Новосибирск

Вульфин Алексей Михайлович, д-р техн. наук, доцент, УГАТУ, г. Уфа

Иванов Андрей Валерьевич, канд. техн. наук, доцент, НГТУ, г. Новосибирск

Ивашук Ольга Александровна, д-р техн. наук, проф., НИУ «БелГУ», г. Белгород

Калимолдаев Махсат Нурадилович, академик НАН РК, д-р физ.-мат. наук, проф.,

РГП на ПХВ «Институт информационных и вычислительных технологий»

КН МНВО РК, Республика Казахстан

Картак Вадим Михайлович, д-р техн. наук, проф., УУНиТ, г. Уфа

Кулаков Станислав Матвеевич, д-р техн. наук, проф., СибГИУ, г. Новокузнецк

Магазев Алексей Анатольевич, д-р физ.-мат. наук, проф., ОмГТУ, г. Омск

Марченко Михаил Александрович, д-р физ.-мат. наук, проф., ИВМиМГ, г. Новосибирск

Мэри Анита Е. А. (Mary Anita E. A.), PhD, Professor, Христианский Университет,

г. Бангалор, Индия

Орлов Сергей Павлович, д-р техн. наук, проф., СамГТУ, г. Самара

Петрунин Юрий Юрьевич, д-р филос. наук, проф., МГУ им. М. В. Ломоносова, г. Москва

Пракаша Г. С. (Prakasha G. S.), PhD, Associate Professor, Христианский Университет, г. Бангалор, Индия

Смирнов Сергей Николаевич, д-р техн. наук, проф., академик Академии криптографии РФ, г. Москва

Сулаво Алексей Евгеньевич, д-р техн. наук, доцент, ОмГТУ, г. Омск

Усатова Ольга Александровна, PhD, Associate Professor, РГП на ПХВ

«Институт информационных и вычислительных технологий» КН МНВО РК, Республика Казахстан

Ходашинский Илья Александрович, д-р техн. наук, проф., ТУСУР, г. Томск

Редакция

Главный редактор

Ложников Павел Сергеевич, д-р техн. наук, доцент, НГТУ, г. Новосибирск

Заместитель главного редактора

Вострецов Алексей Геннадьевич, д-р техн. наук, проф., засл. деятель науки РФ, НГТУ, г. Новосибирск

Заведующий редакцией

Архипова Анастасия Борисовна, канд. техн. наук, доцент, НГТУ, г. Новосибирск

***Журнал зарегистрирован 01.03.2021 Федеральной службой по надзору
в сфере связи, информационных технологий и массовых коммуникаций.
Свидетельство о регистрации ПИ № ФС 77-80320***

Адрес издателя и редакции: 630073, г. Новосибирск, пр. К. Маркса, 20.

E-mail: office@publish.nstu.ru и digital-tech-security@mail.ru

Web site: <http://publish.nstu.ru> и <http://journals.nstu.ru/digital-tech-security/>

Publisher and editorial office adress: 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation

До номера 1 (100) 2021 г. включительно журнал выходил под названием
«Сборник научных трудов НГТУ» (ISSN 2307-6879)

16+

© Коллектив авторов, 2026

© Новосибирский государственный
технический университет, 2026

БЕЗОПАСНОСТЬ ЦИФРОВЫХ ТЕХНОЛОГИЙ

ISSN 2782-2230

№ 3 (122)

2026

СОДЕРЖАНИЕ

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

Кузлякин Д.С., Огнев И.А. Комбинированный анализ ТСП-активности для обнаружения скрытых и распределенных сканирований портов	9
Сильченко Н.С., Пермяков Р.А. Защита каналов связи цифровых двойников от атак в распределенных вычислительных системах .	24
Котов Ю.А., Башкиров А.П. Реализация криптоанализа шифра простой замены с использованием знаковых триграмм.....	41
Синюшкин Т.А., Исаев Г.А. Разработка программного обеспечения для обезличивания персональных данных в ИСПДн Bitrix24.....	64
Гладышев С.А., Голубь С.А. Разработка модели процесса управления инфраструктурными уязвимостями с учетом их приоритета.....	84
Правила для авторов	99

Выпускающий редактор *И.П. Брованова*
Корректор *Л.Н. Кинит*
Компьютерная верстка *С.И. Ткачева*

Лицензия № ИД 04303 от 20.03.01. Подписано в печать 21.09.2026. Выход в свет 25.09.2026
Формат 60×84 1/16. Бумага офсетная. Тираж 300 экз. Уч.-изд. л. 6,04
Печ. л. 6,5. Изд. № 135. Заказ № 177. Цена свободная

Отпечатано в типографии
Новосибирского государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20

Editorial board

Novosibirsk State Technical University

Editorial council

Chairman of the editorial council

Lozhnikov P.S., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Deputy chairman

Belim S.V., Dr. Sc. (Phys. & Math.), Omsk State Technical University, Omsk, RF

Vostretsov A.G., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Kotenko I.V., Dr. Sc. (Eng.), St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, RF

The members of the editorial council

Avdeenko T.V., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Alguliyev R.M.o., Dr. Sc. (Eng.), Azerbaijan National Academy of Sciences, Institute of Information Technology, Baku, AZE

Aleksandrova E.B., Dr. Sc. (Eng.), Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, RF

Anikin I.V., Dr. Sc. (Eng.), Kazan National Research Technical University named after A.N. Tupolev – KAI, Kazan, RF

Haroutunian M.E., Dr. Sc. (Phys. & Math.), Institute for Informatics and Automation Problems of NAS RA, Yerevan, ARM

Babenko M.G., Dr. Sc. (Phys. & Math.), North-Caucasus Federal University, Stavropol, RF

Barankova I.I., Dr. Sc. (Eng.), Magnitogorsk State Technical University, Magnitogorsk, RF

Bezzateev S.V., Dr. Sc. (Eng.), Saint Petersburg State University of Aerospace Instrumentation, St. Petersburg, RF

Vasil'ev V.I., Dr. Sc. (Eng.), Ufa State Aviation Technical University, UFA, RF

Voevoda A.A., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Vostretsov A.G., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Vulfin A.M., Dr. Sc. (Eng.), Ufa University of Science and Technology, UFA, RF

Ivanov A.V., Cand. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Ivashhuk O.A., Dr. Sc. (Eng.), Belgorod State National Research University, Belgorod, RF

Kalimoldayev M.N., Academician NAS RK, Dr. Sc. (Phys. & Math.), RSE on the REU «Institute of information and computational technologies» CS of the MSHE of the RK, RK

Kartak V.M., Dr. Sc. (Eng.), Ufa University of Science and Technology, UFA, RF

Kulakov S.M., Dr. Sc. (Eng.), Siberian State Industrial University, Novokuznetsk, RF

Magazev A.A., Dr. Sc. (Phys. & Math.), Omsk State Technical University, Omsk, RF

Marchenko M.M., Dr. Sc. (Phys. & Math.), The Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk, RF

Mary Anita E.A., PhD, Professor, Christ University, Bangaluru, India

Orlov S.P., Dr. Sc. (Eng.), Samara State technical university, Samara, RF

Petrinin Yu.Yu., Dr. Sc. (Philos.), Lomonosov Moscow State University, Moscow, RF

Prakasha G.S., PhD, Christ University, Bangaluru, India
Smirnov S.N., Dr. Sc. (Eng.), Academy of Cryptography, Moscow, RF
Sulavko A.E., Dr. Sc. (Eng.), Omsk State Technical University, Omsk, RF
Ussatova O.A., PhD, Associate Professor, RSE on the REU “Institute of information and computational technologies” CS of the MSHE of the RK, RK
Hodashinskij I.A., Dr. Sc. (Eng.), Tomsk State University of Control Systems and Radioelectronics, Tomsk, RF

Editorial office

Chief editor

Lozhnikov P.S., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Deputy chief editor

Vostretsov A.G., Dr. Sc. (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Head of the editorial office

Arhipova A.B., Candidate of Science (Eng.), Novosibirsk State Technical University, Novosibirsk, RF

Publisher and editorial adress: 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation

E-mail: office@publish.nstu.ru, digital-tech-security@mail.ru

Web site: <http://publish.nstu.ru>, <http://journals.nstu.ru/digital-tech-security/>

© Authors, 2026

© Novosibirsk State

Technical University, 2026

DIGITAL TECHNOLOGY SECURITY

ISSN 2782-2230

№ 3 (122)

2026

CONTENTS

METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY

Kuzlyakin D.S., Ognev I.A. Combined analysis of TCP activity for detecting stealth and distributed port scanning	9
Silchenko N.S., Permyakov R.A. Protection of digital twin communication channels from attacks in distributed computing systems.....	24
Kotov Yu.A., Bashkirov A.P. Implementation of cryptanalysis of a simple substitution cipher using sign-based trigrams.....	41
Sinyushkin T.A., Isaev G.A. Development of software for personal data anonymization in the Bitrix24 personal data information system.....	64
Gladyshev S.A., Golub S.A. Development of a process model for managing infrastructure vulnerabilities with regard to their priority	84
Rules for authors.....	99

Publishing Editor *I.P. Brovanova*
Editor *L.N. Kinsht*
Computer imposition *S.I. Tkacheva*

License № ID 04303 from 20.03.01. Signed in print September 21, 2026
Date of publication September 25, 2026. Format $60 \times 84 \frac{1}{16}$
Offset Paper. Circulation is 300 copies. Educational-ed. liter. 6,04. printed pages 6,5
Publishing number 135. Order number 177

It is printed in printing house of Novosibirsk State Technical University
630073, Novosibirsk, 20 K. Marx Prospekt

*МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ*

УДК 004

DOI: 10.17212/2782-2230-2026-3-9-23

**КОМБИНИРОВАННЫЙ АНАЛИЗ ТСП-АКТИВНОСТИ
ДЛЯ ОБНАРУЖЕНИЯ СКРЫТЫХ И РАСПРЕДЕЛЕННЫХ
СКАНИРОВАНИЙ ПОРТОВ***

Д.С. КУЗЛЯКИН¹, И.А. ОГНЕВ²

¹ РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, лаборант кафедры защиты информации. E-mail: kuzlyakin.2022@corp.nstu.ru

² РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, старший преподаватель кафедры защиты информации. E-mail: i.ognev@corp.nstu.ru

В статье рассматриваются методы анализа ТСП-активности, предназначенные для обнаружения медленных и распределенных форм сканирования портов, недостаточно эффективно выявляемых традиционными сигнатурными системами. Предложен комбинированный подход, включающий эвристический анализ ТСП-флагов, статистическую оценку интенсивности и энтропии сетевого трафика, графовый анализ взаимодействий между узлами, а также использование структуры Count-Min Sketch для компактного хранения статистики. Представлена методика вычисления интегральной оценки подозрительности (score) на основе совокупности признаков. Экспериментально показано, что предложенные методы обеспечивают точность обнаружения медленного сканирования с интервалом 10 с на уровне 69,5 %, а распределенного сканирования от 20 источников – 73,8 %, что на 17...27 % выше, чем у классических сигнатурных IDS. При этом использование Count-Min Sketch позволяет сократить потребление оперативной памяти на 32...37 % по сравнению с полным хранением статистики портов.

Ключевые слова: ТСП-сканирование, обнаружение аномалий, эвристический анализ, статистический анализ, графовый анализ, Count-Min Sketch, медленное сканирование, распределенное сканирование

ВВЕДЕНИЕ

Сканирование портов является первым и обязательным этапом большинства сетевых атак. Злоумышленник получает информацию об открытых портах, активных сервисах и версиях программного обеспечения, что позволяет

* Статья получена 29 июня 2026 г.

ему подобрать эксплойты и спланировать дальнейшие действия. Современные методы сканирования эволюционировали в сторону повышенной скрытности: используются нестандартные комбинации TCP-флагов (FIN, NULL, XMAS), снижение интенсивности запросов (slow scan) [1–3] и распределение активности между несколькими источниками (distributed scan) [4–6].

Традиционные системы обнаружения вторжений (IDS), основанные на сигнатурном анализе, эффективно выявляют высокоинтенсивные сканирования (например, SYN-сканирование с Nmap по умолчанию), но обладают рядом ограничений. Во-первых, они требуют заранее известных сигнатур атак, что делает их беспомощными против модифицированных или новых методов. Во-вторых, пороговые механизмы не способны обнаружить активность, распределенную во времени или между многими источниками, поскольку отдельные события не превышают установленных пределов.

В связи с этим актуальной является разработка методов анализа TCP-активности, которые не ограничиваются поиском точных совпадений, а выявляют отклонения от нормального поведения сети на основе статистических, временных и пространственных характеристик.

Целью работы стало исследование и разработка методов комбинированного анализа TCP-активности для обнаружения скрытых и распределенных сканирований портов, а также их экспериментальная оценка.

Для достижения поставленной цели были выделены следующие задачи:

- разработать эвристические правила анализа TCP-флагов для выявления нестандартных типов сканирования (FIN, NULL, XMAS);
- реализовать статистические методы оценки интенсивности в скользящем окне и энтропии распределения портов для обнаружения медленных атак;
- создать графовую модель взаимодействия между источниками и портами для выявления распределенных сканирований;
- применить вероятностную структуру Count-Min Sketch для компактного хранения статистики портов с целью снижения потребления памяти;
- провести экспериментальное сравнение предложенного подхода с существующими сигнатурными и потоковыми IDS (Snort, Suricata).

1. ФОРМАЛИЗАЦИЯ ЗАДАЧИ ОБНАРУЖЕНИЯ АНОМАЛИЙ ДЛЯ МОДУЛЯ АНАЛИЗА TCP-АКТИВНОСТИ

Модуль анализа TCP-активности является ключевым компонентом разрабатываемой системы обнаружения. Его входные данные – сетевые пакеты, прошедшие предварительную фильтрацию (только TCP-пакеты, адресован-

ные защищаемому узлу, исключая трафик из whitelist). Для каждого пакета извлекаются IP-адрес источника, порт назначения, TCP-флаги, временная метка, TTL.

Задача модуля – на основе накопленной статистики по каждому источнику определить, является ли его активность легитимной или представляет собой попытку TCP-сканирования. При этом должны выявляться следующие виды аномалий:

- высокоинтенсивное сканирование (SYN, Connect) – большое количество запросов к разным портам за короткое время;
- скрытое сканирование с нестандартными флагами (FIN, NULL, XMAS);
- медленное сканирование – низкая интенсивность, растянутая во времени (интервалы от нескольких секунд до десятков секунд);
- распределенное сканирование – скоординированная активность нескольких источников, каждый из которых генерирует малый объем трафика.

Для решения этой задачи предложен комбинированный подход, объединяющий четыре группы методов: эвристический анализ TCP-флагов, статистический анализ интенсивности и энтропии, графовый анализ взаимодействий и интегральную score-оценку [8–10].

2. МЕТОДЫ АНАЛИЗА TCP-АКТИВНОСТИ

2.1. ЭВРИСТИЧЕСКИЙ АНАЛИЗ TCP-ФЛАГОВ

Эвристический анализ основан на наблюдении, что различные методы TCP-сканирования используют специфические комбинации флагов, редко встречающиеся в легитимном трафике. Модуль включает для каждого источника следующие счетчики за всё время наблюдения [11, 12] (с периодическим сбросом после обнаружения атаки):

- `syn` – количество пакетов с флагом SYN без ACK;
- `ack` – количество пакетов SYN+ACK;
- `fin` – количество пакетов с FIN без SYN;
- `null_scan` – количество пакетов без установленных флагов;
- `xmas` – количество пакетов с комбинацией FIN+PSH+URG;
- `port_count` – количество уникальных портов, к которым были запросы.

Эвристические правила (пороги настраиваются, но типичные значения приведены ниже):

- SYN-сканирование: если `syn > 100` и `port_count > 20`, фиксируется атака NMAP_SYN;

- FIN-сканирование: если $\text{fin} > 80$ и $\text{port_count} > 10 - \text{FIN_SCAN}$;
- NULL-сканирование: $\text{null_scan} > 30$ и $\text{port_count} > 10 - \text{NULL_SCAN}$;
- XMAS-сканирование: $\text{xmas} > 30$ и $\text{port_count} > 10 - \text{XMAS_SCAN}$;
- Connect-сканирование: $\text{syn} > 50$ и $\text{ack} > 0.5 * \text{syn}$ и $\text{port_count} > 10 - \text{CONNECT_SCAN}$.

Эти правила эффективны для интенсивных атак, но не работают для медленных и распределенных вариантов, где абсолютные значения счетчиков могут оставаться низкими. Поэтому дополнительно применяются статистические и графовые методы.

2.2. СТАТИСТИЧЕСКИЙ АНАЛИЗ ИНТЕНСИВНОСТИ И ВРЕМЕННЫХ ПАТТЕРНОВ

2.2.1. АНАЛИЗ ИНТЕНСИВНОСТИ В СКОЛЬЗЯЩЕМ ОКНЕ

Для каждого источника поддерживается временное окно длительностью 10 секунд, в котором подсчитывается количество SYN-пакетов (функция `window_rate_us`). Это позволяет выявлять всплески активности. Правило: если $\text{window_rate} > 50$ и $\text{port_count} > 20$, добавляется признак BURST. Такая ситуация характерна для быстрого сканирования или DoS-атаки.

2.2.2. АНАЛИЗ СРЕДНЕГО ИНТЕРВАЛА И ДЖИТТЕРА

Для обнаружения медленного сканирования, при котором интервалы между пакетами составляют несколько секунд, накапливаются статистики:

- средний интервал `avg_interval` (экспоненциально взвешенное скользящее среднее с коэффициентом 0,8)

$$\bar{t}_n = 0,8\bar{t}_{n-1} + 0,2\Delta t_n;$$

- джиттер `jitter` – накопленное абсолютное отклонение

$$J_n = J_{n-1} + |\Delta t_n - \bar{t}_{n-1}|.$$

Если $\text{avg_interval} < 0,05$ с (т. е. более 20 пакетов в секунду) и $\text{port_count} > 20$, добавляется признак TIMING. При малых интервалах это указывает на высокую частоту – признак интенсивного сканирования.

Для медленного сканирования используется другой критерий: если $\text{port_count} > \text{slow_scan_ports}$ (например, 15) и длительность наблюдения $\text{duration} > \text{slow_scan_duration}$ (300 секунд), то добавляется признак SLOW

с дополнительными баллами при длительности более 5 или 10 минут. Этот критерий не требует высокой мгновенной интенсивности и позволяет обнаруживать атаки, растянутые во времени.

2.2.3. ЭНТРОПИЙНЫЙ АНАЛИЗ РАСПРЕДЕЛЕНИЯ ПОРТОВ

При сканировании большого диапазона портов распределение запросов по портам становится близким к равномерному, в то время как легитимный трафик обычно концентрируется на небольшом числе портов (80, 443, 22 и т. д.). Для количественной оценки этого свойства используется энтропия Шеннона [13–15]:

$$H = - \sum_{i=1}^N p_i \log_2 p_i .$$

где p_i – оценка вероятности обращения к порту i . Для вычисления частот используется структура Count-Min Sketch (см. раздел 2.4), что позволяет оценить энтропию без хранения полного вектора частот. При $H > 4,0$ и $\text{port_count} > 20$ фиксируется признак ENTROPY.

2.3. ГРАФОВЫЙ АНАЛИЗ РАСПРЕДЕЛЕННОГО СКАНИРОВАНИЯ

Распределенное сканирование характеризуется тем, что много разных источников обращаются к одним и тем же портам (или к разным, но с высокой степенью «новизны» для каждого источника). Для выявления таких паттернов строится граф, вершинами которого являются IP-адреса источников и порты назначения, а ребра соответствуют факту обращения источника к порту [16–18].

В модуле реализованы следующие структуры:

- SrcNode – для каждого источника хранится буфер времен последних обращений и буфер времени первого обращения к новым портам;
- PortNode – для каждого порта хранится количество уникальных источников, обратившихся к нему;
- GraphEdge – ребро (источник, порт) с временем первого и последнего обращения.

Периодически (каждые 10 секунд) запускается анализ графа за временное окно (по умолчанию 30 секунд). Для каждого источника вычисляются:

- new_ports – количество портов, к которым источник обратился впервые (т. е. ранее этого окна обращения к данным портам не было);

- `total_probes` – общее число обращений;
- `repeat_ratio = 1 - new_ports / total_probes` – доля повторных обращений.

Если `new_ports` $\geq$ `fanout_threshold` (например, 20) и `repeat_ratio` $<$ `min_repeat_ratio` (0,3), источник считается участником распределенного сканирования. Логика обоснования: при распределенной разведке каждый источник обычно сканирует новые порты (низкий `repeat_ratio`) в большом количестве (высокий `fanout`). Дополнительно анализируется глобальная метрика – суммарное значение `new_ports` по всем источникам. Если она превышает `global_fanout_alert` (100), фиксируется координированная распределенная кампания.

2.4. ВЕРОЯТНОСТНЫЕ СТРУКТУРЫ ДАННЫХ COUNT-MIN SKETCH

Хранение точной частоты обращений к каждому из 65535 портов потребовало бы значительного объема памяти (например, массив `uint32_t` на 65535 элементов – 256 кБ на один источник, что неприемлемо при многих источниках). Для решения этой проблемы используется Count-Min Sketch – вероятностная структура данных, позволяющая оценивать частоту элементов с фиксированной памятью [19–21].

Count-Min Sketch представляет собой матрицу счетчиков размером `depth` $\times$ `width` (в реализации `depth = 4`, `width = 1024`). Каждый элемент хешируется несколькими независимыми хеш-функциями, и для каждой строки инкрементируется соответствующий счетчик. Оценка частоты элемента равна минимальному значению среди счетчиков по всем хешам.

В модуле Count-Min Sketch используется для оценки частоты обращения к портам при энтропии (метод `calc_entropy`) и быстром обновлении статистики без выделения памяти под новые порты.

Память, занимаемая одним sketch: $4 \times 1024 \times 4$ байта = 16 кБ. Это в 16 раз меньше точного массива. Эксперименты показали, что при сканировании до 5000 портов погрешность оценки не превышает 5 %, что достаточно для энтропийного анализа.

2.5. ИНТЕГРАЛЬНАЯ SCORE-ОЦЕНКА

Для комплексной оценки подозрительности сетевой активности разработана интегральная score-оценка, агрегирующая все описанные выше признаки. Каждый признак вносит вклад в итоговый балл в соответствии со своей значимостью, определенной экспертным путем и подтвержденной экспериментальными данными. Ниже приведены веса признаков и условия их срабатывания [8–10].

Для наглядности все признаки, их весовые коэффициенты и пороговые условия сведены в табл. 1. Выбор весов обусловлен различной информативностью признаков: наибольший вес (30 баллов) присвоен признаку BURST, характеризующему резкие всплески активности, а наименьший (10 баллов) – признаку LINEAR, указывающему на последовательный перебор портов. Промежуточные значения весов отражают вклад каждого признака в общую оценку подозрительности.

Таблица 1

Table 1

Вес и значимость признака
Weight and significance of the feature

Признак	Вес	Условие
BURST	30	rate > 50 и port_count > 20
FAST	25	port_count > fast_ports (50), rate > fast_rate (35), duration < fast_duration (90 с)
TIMING	15	avg_interval < 0,05 и port_count > 20
LINEAR	10	seq_score > 0,4 (последовательный перебор портов)
ENTROPY	20	H > 4,0 и port_count > 20
SLOW	25	port_count > slow_ports (15) и duration > slow_duration (300 с), дополнительные +10 при duration > 600 с, +5 при duration > 300 с
Nmap сигнатура	20	сработало одно из эвристических правил из раздела 2.1

Итоговое значение score суммируется. Затем в зависимости от его значения формируется уровень угрозы:

- score $\geq$ 60 – CRITICAL SCAN;
- score $\geq$ 40 – HIGH RISK SCAN;
- score $\geq$ 20 – SUSPICIOUS SCAN.

Кроме того, значение score влияет на динамическую настройку времени жизни записи об источнике (ttl_seconds): для критических атак TTL увеличивается до 600 секунд (чтобы не потерять информацию об активном злоумышленнике), для обычных источников – 60 секунд.

После срабатывания атаки (формирования alert) статистика источника частично сбрасывается: счетчики syn, ack, fin и т. д. умножаются на 0,8 (экспо-

ненциальное забывание), но port_count и битовая карта портов полностью обнуляются. Это предотвращает повторные срабатывания по той же атаке и позволяет адаптироваться к долговременной активности.

3. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА МЕТОДОВ

3.1. КОНФИГУРАЦИЯ ЭКСПЕРИМЕНТА

Тестирование проводилось на MacBook Pro M1 Pro (10 ядер CPU, 16 ГБ ОЗУ). Генерация атак выполнялась с помощью Nmap и самописных скриптов [12]. Фоновый легитимный трафик включал веб-серфинг, DNS, SSH. Основные сценарии:

- SYN сканирование портов 1...5000 со скоростью 800...4000 пакетов в секунду (30 запусков);
- FIN-, NULL-, XMAS-сканирование (по 20 запусков);
- медленное сканирование с интервалами 1, 3, 5, 10 секунд (по 10 запусков);
- распределенное сканирование от 5 до 20 источников (по 10 запусков).

3.2. РЕЗУЛЬТАТЫ ОБНАРУЖЕНИЯ

Для различных типов сканирования допускается точность обнаружения (табл. 2).

Таблица 2

Table 2

Точность обнаружения различных типов сканирования

Detection accuracy of different scan types

Тип сканирования	Accuracy, %	Precision, %	Recall, %
SYN	96,2	96,0	96,4
FIN	95,2	94,8	95,6
NULL	94,1	93,5	94,7
XMAS	95,8	96,0	95,6
Slow scan	69,5	71,0	68,0
Distributed	73,8	75,5	72,0

Для интенсивных атак методы показали высокую эффективность (94...96 %). Медленное сканирование с интервалом 1 секунда обнаруживается почти так же хорошо, как интенсивное, поскольку окно в 10 секунд всё еще

фиксирует достаточное количество пакетов. При интервале 10 секунд точность падает до 69,5 %. Основная сложность связана с тем, что за 10-секундное окно может не набраться ни одного пакета, и требуются долговременные накопления (признак SLOW), которые могут давать ложные срабатывания на редком легитимном трафике.

Распределенное сканирование с пятью источниками обнаруживается очень хорошо (94 %), но с ростом числа источников точность снижается. При 20 источниках каждый в среднем сканирует лишь 100 портов (из 3000 целевых), что приближается к порогу `fanout_threshold = 20`, и некоторые источники могут не превысить его. Тем не менее глобальный критерий `global_fanout_alert` позволяет выявить саму кампанию.

3.3. ЭФФЕКТИВНОСТЬ ИСПОЛЬЗОВАНИЯ COUNT-MIN SKETCH

Сравнение потребления памяти для хранения статистики портов на один источник приведено в табл. 2. Точное хранение (битовая карта + массив частот) требует более 512 кБ на источник. Использование CMS снижает этот объем до 16 кБ (в 32 раза). При 1000 активных источников (что бывает при DDoS-атаках или сканировании из ботнета) экономия памяти составляет десятки мегабайт (табл. 3).

Таблица 3

Table 3

Сравнение памяти для хранения статистики портов

Comparison of memory for storing port statistics

Метод	Память на источник	Память на 1000 источников
Битовая карта (65 536 бит)	8 кБ	8 МБ
Точные счетчики (uint32 на порт)	256 кБ	256 МБ
Count-Min Sketch (4×1024×4)	16 кБ	16 МБ

4. СРАВНЕНИЕ С СУЩЕСТВУЮЩИМИ МЕТОДАМИ

Сравнение предложенного комбинированного подхода с классическими сигнатурными IDS (Snort) и потоковой Suricata представлено в табл. 4. Данные для Snort и Suricata получены при их настройке по умолчанию с подключенными правилами для обнаружения сканирования.

Т а б л и ц а 4

T a b l e 4

Сравнение точности обнаружения скрытых атак
Comparison of the accuracy of detection of hidden attacks

Тип атаки	Предложенные методы	Snort	Suricata
Slow scan	69,5 %	52,4 %	70,1 %
Distributed scan	73,8 %	46,8 %	67,4 %

Программное обеспечение Snort, основанное на сигнатурах, не способно накапливать статистику за длительные интервалы и выявлять скоординированную активность разных источников, поэтому его эффективность низка; Suricata имеет механизмы потокового анализа, что позволяет ему обнаруживать медленное сканирование наравне с нашей системой (разница в пределах погрешности). Однако для распределенного сканирования Suricata уступает 6,4 %, поскольку его корреляционные механизмы менее развиты.

По вычислительной эффективности предложенные методы также выигрывают: загрузка CPU при обработке 25 000 пакетов в секунду составила 18 % (против 36 % у Snort и 29 % у Suricata). Это объясняется отказом от глубокого анализа пакетов и использованием легковесных структур данных.

ЗАКЛЮЧЕНИЕ

В статье представлен комплекс методов анализа TCP-активности для обнаружения скрытых и распределенных форм сканирования портов. Основные результаты:

- эвристический анализ TCP-флагов позволяет с высокой точностью (94...96 %) выявлять интенсивные SYN, FIN, NULL, XMAS сканирования;
- статистический анализ интенсивности в скользящем окне и долгосрочных интервалов дает возможность обнаруживать медленное сканирование. Точность для интервала 10 секунд достигла 69,5 %, что на 17 % выше, чем у Snort;
- графовый анализ взаимодействий между источниками и портами эффективен для выявления распределенных атак: точность 73,8 % для 20 источников, что на 27 % выше, чем у Snort, и на 6 % выше, чем у Suricata;
- использование Count-Min Sketch позволило сократить память для хранения частот портов в 32 раза (до 16 кБ на источник) без значимой потери точности энтропийной оценки;

- интегральная score-оценка объединяет все признаки, обеспечивая баланс между чувствительностью и устойчивостью к ложным срабатываниям (общий FPR 2,4 %).

Предложенные методы могут быть реализованы в виде легковесного модуля, работающего в реальном времени на узлах с ограниченными ресурсами. Дальнейшие исследования направлены на применение методов машинного обучения для автоматического определения пороговых значений и адаптации к профилю нормального трафика конкретной сети.

СПИСОК ЛИТЕРАТУРЫ

1. *Muraleedharan N., Janet B.* Slow TCP port scan detection using flow data // International Conference on Security, Privacy and Data Analytics. – Singapore: Springer, 2022. – P. 17–32.
2. *Miranda D., Monteiro R., Silva J.M.C.* p4SD: A lightweight port scan detection for programmable networks // International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia. – IEEE, 2025. – P. 1–6. – DOI: 10.23919/SoftCOM66362.2025.11197439.
3. *Almseidin M., Al-Kasassbeh M., Kovacs S.* Detecting slow port scan using fuzzy rule interpolation // 2019 2nd International Conference on New Trends in Computing Sciences (ICTCS). – Amman: IEEE, 2019. – P. 1–6.
4. *Ring M., Landes D., Hotho A.* Detection of slow port scans in flow-based network traffic // PLoS one. – 2018. – Vol. 13 (9). – P. e0204507.
5. *Haas S., Wilkens F., Fischer M.* Scan correlation – Revealing distributed scan campaigns // 2020 IEEE/IFIP Network Operations and Management Symposium (NOMS). – 2020. – P. 1–6. – DOI: 10.48550/arXiv.2003.05188.
6. *Griffioen H., Doerr C.* Discovering collaboration: Unveiling slow, distributed scanners based on common header field patterns // 2020 IEEE/IFIP Network Operations and Management Symposium (NOMS). – 2020. – P. 1–9. – DOI: 10.1109/NOMS47738.2020.9110444.
7. *Gates C.* Coordinated scan detection // Proceedings of the Network and Distributed System Security Symposium (NDSS 2009). – San Diego, California, 2009.
8. *Токарев Д.А., Погребан С.В.* Аномалии трафика в IP сетях и их обнаружение // Известия Института инженерной физики. – 2013. – № 4. – С. 13–17.
9. Формализованная модель обнаружения сканирования портов / Е.В. Ананьин, И.С. Кожевникова, А.В. Лысенко, А.В. Никишова // Молодой ученый. – 2016. – № 23 (127). – С. 9–12.
10. *Головачева Н.А., Андрейченко Н.Е.* Исследование методов прогнозирования атак на информационные системы // НБИ технологии. – 2022. – Т. 16, № 3. – С. 11–15.

11. PSAD – Port Scan Attack Detector (man pages) // PSAD Documentation.
12. TCP FIN, NULL, and Xmas Scans // Nmap Network Scanning Documentation.
13. *Ayoub L., Shurman M.* A lightweight entropy-based framework for multi-attack detection in software-defined networks (SDN) // 2025 26th International Arab Conference on Information Technology (ACIT). – IEEE, 2025. – P. 179–183. – DOI: 10.1109/ACIT68900.2025.11510552.
14. *Timcenko V., Gajin S.* Machine learning enhanced entropy-based network anomaly detection // Advances in Electrical and Computer Engineering. – 2021. – Vol. 21 (4). – P. 51–60.
15. Efficient and secure statistical port scan detection scheme / H. Majed, H.N. Noura, O. Salman, A. Chehab, R. Couturier // Mobile, Secure, and Programmable Networking: 6th International Conference, MSPN 2020. – Cham: Springer, 2021. – P. 72–88.
16. A hypergraph-based machine learning ensemble network intrusion detection system / Z.-Z. Lin, T.D. Pike, M.M. Bailey, N.D. Bastian // IEEE Transactions on Systems, Man, and Cybernetics: Systems. – 2024. – Vol. 54 (11). – P. 6911–6923.
17. Leveraging graph-based representations to enhance machine learning performance in IIoT network security and attack detection / B. Alwasel, A. Aldribi, M. Alreshoodi, I.S. Alsukayti, M. Alsuhaibani // Applied Sciences. – 2023. – Vol. 13 (13). – P. 7774.
18. A cooperation-aware virtual network function for proactive detection of distributed port scanning / I.J. Sanz, M.A. Lopez, D.M.F. Mattos, O.C.M.B. Duarte // 2017 1st Cyber Security in Networking Conference (CSNet'17). – Rio de Janeiro, Brazil, 2017. – P. 1–8.
19. *Martínez Aguilar R., Reviriego P., Larrabeiti D.* Combined filtering and frequency estimation with the integrated Xor filter and count-min sketch // 2024 IEEE 25th International Conference on High Performance Switching and Routing (HPSR). – IEEE, 2024. – P. 155–160. – DOI: 10.1109/HPSR62440.2024.10635957.
20. *Ekle O.A., Eberle W.* Adaptive-graph sketch: Real-time edge anomaly detection via multi-layer tensor sketching and temporal decay // 2025 IEEE International Conference on Knowledge Graph (ICKG). – IEEE, 2025. – P. 74–81.
21. PD-CPS: A practical scheme for detecting covert port scans in high-speed networks / H. Wu, Z. Shao, F. Yang, G. Cheng, X. Hu, J. Ren, W. Wang // Computer Networks. – 2023. – Vol. 231. – P. 109825. – DOI: 10.1016/j.comnet.2023.109825.

Кузлякин Дмитрий Сергеевич, лаборант кафедры защиты информации Новосибирского государственного технического университета. E-mail: kuzlyakin.2022@corp.nstu.ru

Огнев Иван Александрович, старший преподаватель кафедры защиты информации Новосибирского государственного технического университета. E-mail: i.ognev@corp.nstu.ru

DOI: 10.17212/2782-2230-2026-3-9-23

Combined analysis of TCP activity for detecting stealth and distributed port scanning*

D.S. Kuzlyakin¹, I.A. Ognev²

¹ *Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, laboratory assistant at the Department of Information Security. E-mail: kuzlyakin.2022@corp.nstu.ru*

² *Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, senior lecturer at the Department of Information Protection. E-mail: i.ognev@corp.nstu.ru*

This article discusses TCP activity analysis methods designed to detect slow and distributed forms of port scanning that are insufficiently detected by traditional signature-based IDSs. A combined approach is proposed, including heuristic analysis of TCP flags, statistical assessment of network traffic intensity and entropy, graph analysis of interactions between nodes, and the use of the Count-Min Sketch structure for compact storage of statistics. A methodology for calculating an integrated suspicion score based on a set of features is presented. Experiments have shown that the proposed methods provide 69.5 % detection accuracy for slow scanning with a 10-s interval and 73.8 % for distributed scanning from 20 sources, which is 17–27 % higher than that of classic signature-based IDSs. At the same time, using Count-Min Sketch allows you to reduce RAM consumption by 32–37 % compared to storing full port statistics.

Keywords: TCP scanning, anomaly detection, heuristic analysis, statistical analysis, graph analysis, Count-Min Sketch, slow scan, distributed scan

REFERENCES

1. Muraleedharan N., Janet B. Slow TCP port scan detection using flow data. *International Conference on Security, Privacy and Data Analytics*, Singapore. Springer, 2022, pp. 17–32.

* Received 29 June 2026.

2. Miranda D., Monteiro R., Silva J.M.C. p4SD: A lightweight port scan detection for programmable networks. *International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, Split, Croatia. IEEE, 2025, pp. 1–6. DOI: 10.23919/SoftCOM66362.2025.11197439.
3. Almseidin M., Al-Kasassbeh M., Kovacs S. Detecting slow port scan using fuzzy rule interpolation. *2019 2nd International Conference on New Trends in Computing Sciences (ICTCS)*. Amman, IEEE, 2019, pp. 1–6.
4. Ring M., Landes D., Hotho A. Detection of slow port scans in flow-based network traffic. *PLoS one*, 2018, vol. 13 (9), p. e0204507.
5. Haas S., Wilkens F., Fischer M. Scan correlation – Revealing distributed scan campaigns. *2020 IEEE/IFIP Network Operations and Management Symposium (NOMS)*, 2020, pp. 1–6. DOI: 10.48550/arXiv.2003.05188.
6. Griffioen H., Doerr C. Discovering collaboration: Unveiling slow, distributed scanners based on common header field patterns. *2020 IEEE/IFIP Network Operations and Management Symposium (NOMS)*, 2020, pp. 1–9. DOI: 10.1109/NOMS47738.2020.9110444.
7. Gates C. Coordinated scan detection. *Proceedings of the Network and Distributed System Security Symposium (NDSS 2009)*, San Diego, California, 2009.
8. Tokarev D.A., Pogreban S.V. Anomalii trafika v IP setyakh i ikh obnaruzhenie [Traffic anomalies in IP networks and their detection]. *Izvestiya Instituta inzhenernoi fiziki*, 2013, no. 4, pp. 13–17. (In Russian).
9. Anan'in E.V., Kozhevnikova I.S., Lysenko A.V., Nikishova A.V. Formalizovannaya model' obnaruzheniya skanirovaniya portov [Formalized model for detecting port scanning]. *Molodoi uchenyi = Young Scientist*, 2016, no. 23, pp. 9–12.
10. Golovacheva N.A., Andreichenko N.E. Issledovanie metodov prognozirovaniya atak na informatsionnye sistemy [Investigation of methods for predicting attacks on information systems]. *NBI tekhnologii = NBI Technologies*, 2022, vol. 16, no. 3, pp. 11–15.
11. PSAD – Port Scan Attack Detector (man pages). *PSAD Documentation*.
12. TCP FIN, NULL, and Xmas Scans. *Nmap Network Scanning Documentation*.
13. Ayoub L., Shurman M. A lightweight entropy-based framework for multi-attack detection in software-defined networks (SDN). *2025 26th International Arab Conference on Information Technology (ACIT)*, 2025, pp. 179–183. DOI: 10.1109/ACIT68900.2025.11510552.
14. Timcenko V., Gajin S. Machine learning enhanced entropy-based network anomaly detection. *Advances in Electrical and Computer Engineering*, 2021, vol. 21 (4), pp. 51–60.
15. Majed H., Noura H.N., Salman O., Chehab A., Couturier R. Efficient and secure statistical port scan detection scheme. *Mobile, Secure, and Programmable*

Networking: 6th International Conference, MSPN 2020. Cham, Springer, 2021, pp. 72–88.

16. Lin Z.-Z., Pike T.D., Bailey M.M., Bastian N.D. A hypergraph-based machine learning ensemble network intrusion detection system. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2024, vol. 54 (11), pp. 6911–6923.

17. Alwasel B., Aldribi A., Alreshoodi M., Alsukayti I.S., Alsuhaibani M. Leveraging graph-based representations to enhance machine learning performance in IIoT network security and attack detection. *Applied Sciences*, 2023, vol. 13 (13), p. 7774.

18. Sanz I.J., Lopez M.A., Mattos D.M.F., Duarte O.C.M.B. A cooperation-aware virtual network function for proactive detection of distributed port scanning. *2017 1st Cyber Security in Networking Conference (CSNet'17)*, Rio de Janeiro, Brazil, 2017, pp. 1–8.

19. Martínez Aguilar R., Reviriego P., Larrabeiti D. Combined filtering and frequency estimation with the integrated Xor filter and count-min sketch. *2024 IEEE 25th International Conference on High Performance Switching and Routing (HPSR)*, 2024, pp. 155–160. DOI: 10.1109/HPSR62440.2024.10635957.

20. Ekle O.A., Eberle W. Adaptive-graphsketch: Real-time edge anomaly detection via multi-layer tensor sketching and temporal decay. *2025 IEEE International Conference on Knowledge Graph (ICKG)*. IEEE, 2025, pp. 74–81.

21. Wu H., Shao Z., Yang F., Cheng G., Hu X., Ren J., Wang W. PD-CPS: A practical scheme for detecting covert port scans in high-speed networks. *Computer Networks*, 2023, vol. 231, p. 109825. DOI: 10.1016/j.comnet.2023.109825.

Для цитирования:

Кузлякин Д.С., Огнев И.А. Комбинированный анализ TCP-активности для обнаружения скрытых и распределенных сканирований портов // Безопасность цифровых технологий. – 2026. – № 3 (122). – С. 9–23. – DOI: 10.17212/2782-2230-2026-3-9-23.

For citation:

Kuzlyakin D.S., Ognev I.A. Kombinirovannyi analiz TCP-aktivnosti dlya obnaruzheniya skrytykh i raspredeleennykh skanirovaniy portov [Combined analysis of TCP activity for detecting stealth and distributed port scanning]. *Bezopasnost' tsifrovyykh tekhnologii = Digital Technology Security*, 2026, no. 3 (122), pp. 9–23. DOI: 10.17212/2782-2230-2026-3-9-23.

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

УДК 004

DOI: 10.17212/2782-2230-2026-3-24-40

**ЗАЩИТА КАНАЛОВ СВЯЗИ ЦИФРОВЫХ ДВОЙНИКОВ
ОТ АТАК В РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ
СИСТЕМАХ***

Н.С. СИЛЬЧЕНКО¹, Р.А. ПЕРМЯКОВ²

¹ РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, лаборант кафедры защиты информации. E-mail: silchenko.2022@stud.nstu.ru

² РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, лаборант кафедры защиты информации E-mail: permyakov@corp.nstu.ru

Технология цифровых двойников активно внедряется в промышленный интернет вещей для мониторинга и управления сложными киберфизическими системами в режиме реального времени. Двусторонний характер обмена данными между физическим объектом и его виртуальной копией порождает дополнительные векторы атак, причем особую опасность представляют атаки на каналы связи, использующие легковесный протокол MQTT. В статье разработана модель угроз безопасности информации для каналов связи цифровых двойников в соответствии с методикой ФСТЭК России, выделено шесть актуальных угроз, включая подмену данных, повторное воспроизведение и отказ в обслуживании. Создан экспериментальный стенд на базе эмулятора датчика температуры, MQTT-брокера HiveMQ и платформы цифровых двойников Eclipse Ditto, развернутых в Docker-контейнерах. Реализованы сценарии атак MITM с подменой данных, Replay-атаки и DoS-атаки. Показано, что незащищенный MQTT-канал уязвим для всех перечисленных угроз. Предложен комплекс механизмов защиты: TLS-шифрование канала, проверка временных меток в JavaScript-маппере и правила iptables для ограничения интенсивности трафика. Экспериментально подтверждена эффективность предложенных мер: MITM-атака становится невозможной при корректной проверке сертификата, Replay-атака предотвращается для рассматриваемого сценария отбрасыванием устаревших сообщений, DoS-атака существенно снижается за счет фильтрации на уровне межсетевого экрана.

Ключевые слова: цифровой двойник, MQTT, TLS, MITM, Replay-атака, DoS-атака, iptables, Eclipse Ditto, HiveMQ, распределенные вычислительные системы, модель угроз

* Статья получена 02 июля 2026 г.

ВВЕДЕНИЕ

В условиях цифровой трансформации промышленности и перехода к концепции Industry 4.0 технология цифровых двойников (ЦД) становится ключевым инструментом для мониторинга, анализа и управления сложными техническими системами в реальном времени. Цифровой двойник представляет собой виртуальную копию физического объекта, синхронизированную с ним по двунаправленным каналам связи [1, 2]. Эта синхронизация позволяет не только наблюдать за состоянием объекта, но и прогнозировать его поведение, выявлять аномалии и даже формировать управляющие воздействия.

Однако именно двусторонний характер обмена данными существенно расширяет поверхность атаки. Злоумышленник, получивший доступ к каналу связи, может не только перехватывать конфиденциальную телеметрию, но и модифицировать ее в реальном времени, повторно воспроизводить легитимные сообщения или инициировать отказ в обслуживании [3, 4]. Особую опасность представляет протокол MQTT, широко применяемый в промышленном интернете вещей благодаря своей легкости и поддержке модели публикации/подписки. Стандартная конфигурация MQTT не предусматривает встроенного шифрования, аутентификации или контроля целостности, что делает протокол уязвимым для атак типа «человек посередине», повторного воспроизведения и отказа в обслуживании [5, 6].

Существующие подходы к обеспечению информационной безопасности, например построение центров мониторинга и реагирования на инциденты (SOC), ориентированы на обнаружение уже известных паттернов атак и часто являются реактивными. Цифровой двойник, напротив, благодаря своей способности моделировать будущие состояния системы может реализовать проактивную парадигму защиты, выявляя аномалии на ранних стадиях [7]. Однако для этого сам канал связи, по которому передаются данные, должен быть надежно защищен.

Цель настоящей работы – исследование и экспериментальная оценка механизмов защиты каналов связи цифровых двойников от атак в распределенных вычислительных системах. В работе решаются следующие задачи: анализ архитектуры каналов связи и разработка модели угроз по методике ФСТЭК России; создание экспериментального стенда, эмулирующего работу цифрового двойника на базе протокола MQTT; реализация сценариев типовых атак; внедрение комплекса защитных механизмов и оценка их эффективности, включая влияние на временные характеристики передачи данных.

1. АРХИТЕКТУРА КАНАЛОВ СВЯЗИ ЦИФРОВЫХ ДВОЙНИКОВ И МОДЕЛЬ УГРОЗ

Ключевой особенностью цифрового двойника является двусторонняя автоматическая связь с физическим объектом. По каналу «объект → двойник» передаются показания датчиков, параметры работы оборудования, журналы событий. По обратному каналу «двойник → объект» могут отправляться команды управления, рекомендации, аварийные сигналы [1, 8].

В качестве транспорта в распределенных системах часто используется протокол MQTT. Он работает поверх TCP, поддерживает три уровня качества обслуживания (QoS 0,1,2) и является легковесным, что важно для устройств с ограниченными ресурсами. Однако стандартная реализация MQTT не содержит средств шифрования, аутентификации и контроля целостности сообщений.

На основе методики ФСТЭК России [9] и банка данных угроз [10] была разработана модель угроз для экспериментального стенда, включающего эмулятор датчика, MQTT-брокер и платформу цифровых двойников. Выделены следующие актуальные угрозы (табл. 1).

Таблица 1

Table 1

Актуальные угрозы безопасности информации

Current threats to information security

Идентификатор	Наименование угрозы	Краткое описание
УБИ.069	Неправомерные действия в каналах связи	MITM-атака с подменой значения температуры
УБИ.006	Внедрение кода или данных	Инъекция ложных телеметрических сообщений
УБИ.083	Несанкционированный доступ	Подключение к MQTT-брокеру без аутентификации
УБИ.140	Отказ в обслуживании	DoS-атака путём флуда сообщениями
УБИ.192	Использование уязвимых версий ПО	Эксплуатация известных уязвимостей в HiveMQ, Ditto
УБИ.116	Перехват данных в сети	Сниффинг MQTT-трафика

Основными уязвимостями стенда являются: передача данных в открытом виде, отсутствие аутентификации, отсутствие проверки целостности и свежести сообщений, а также отсутствие rate limiting на уровне брокера.

2. ЭКСПЕРИМЕНТАЛЬНЫЙ СТЕНД

Для экспериментального исследования был разработан стенд на базе двух виртуальных машин с ОС Ubuntu 22.04 LTS и одной машины с ОС Windows 10. Серверные компоненты (MQTT-брокер HiveMQ и платформа Eclipse Ditto) развернуты в Docker-контейнерах (табл. 2).

Т а б л и ц а 2
T a b l e 2

Конфигурация узлов стенда
Configuration of the stand nodes

Роль узла	ОС	CPU	RAM	IP-адрес
Сервер приложений (брокер + Ditto)	Ubuntu 22.04 LTS	4 vCPU	8 GB	192.168.0.156
Атакующая машина (MITM, DoS, Replay)	Ubuntu 22.04 LTS	2 vCPU	4 GB	192.168.0.192
Эмулятор датчика	Windows 10 Pro	4 vCPU	4 GB	192.168.0.155

Эмулятор датчика температуры написан на Python с использованием библиотеки paho-mqtt. Каждое сообщение содержит JSON-структуру, представленную на рис. 1.

```
{
  "sensor_id": "temp-sensor-91",
  "temperature": 20.1,
  "timestamp": 1776144037.2124453
}
```

Рис. 1. JSON-структура телеметрического сообщения от сенсора

Fig. 1. JSON structure of the telemetry message from the sensor

Цифровой двойник создан в Eclipse Ditto как Thing sensor:temp1 с feature temperature, содержащей свойства value и timestamp. Интеграция с MQTT реализована через встроенный MQTT-коннектор, который подписывается на топик sensors/temperature. Для преобразования входящих сообщений во внутренний протокол Ditto используется JavaScript-маппер. Такая архитектура обеспечивает гибкость экспериментальной среды и позволяет в реальном времени отслеживать изменения состояния цифрового двойника при поступлении телеметрии.

Схема взаимодействия компонентов представлена на рис. 2 и включает следующие потоки:

- штатная передача данных от эмулятора к брокеру (порт 1883 без шифрования, порт 8883 – с TLS);
- в режиме MITM-атаки трафик перенаправляется на прокси (порт 1884), который модифицирует значение температуры;
- DoS-атака генерирует многопоточный флуд сообщениями;
- Replay-атака состоит из фазы sniffинга (запись сообщений) и фазы воспроизведения.

Такое разделение потоков позволяет последовательно эмулировать каждый тип угроз и оценивать эффективность соответствующих защитных механизмов в контролируемых условиях.

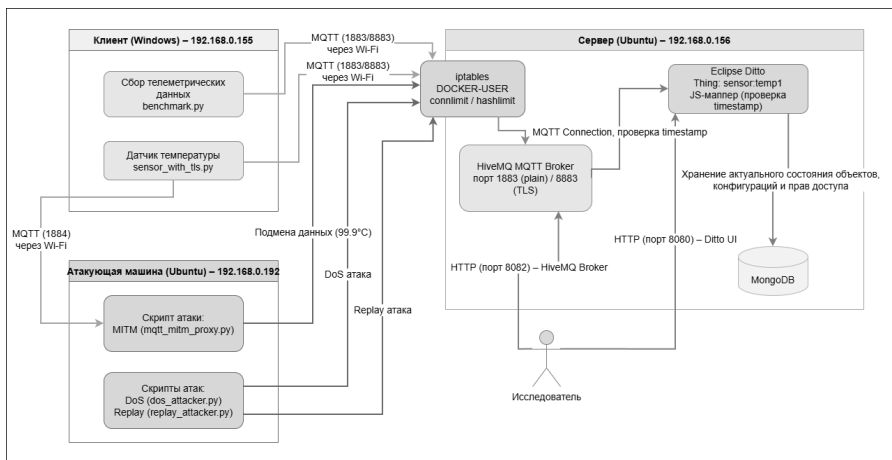


Рис. 2. Схема взаимодействия компонентов

Fig. 2. Component interaction diagram

3. МОДЕЛИРОВАНИЕ АТАК

3.1. MITM-АТАКА С ПОДМЕНОЙ ДАННЫХ

Разработан прокси-сервер на Python, который перехватывает TCP-соединение между эмулятором и брокером. Прокси-сервер анализирует проходящие MQTT-сообщения, находит в JSON-полезной нагрузке поле temperature и заменяет его значение на 99,9 °C. В незащищенном канале (порт 1883) атака успешна: цифровой двойник отображает ложное значение температуры. При этом в логах прокси-сервера фиксируются сообщения MODIFIED (рис. 3). Успех атаки ожидаемый и объясняется отсутствием шифрования. Однако важно отметить, что даже при TLS атака MITM возможна, если клиент не проверяет сертификат, – это ограничение следует учитывать.

Incoming Thing Updates

Reset

Add basic filter...

Revision	Action	Path	Field(s)	Modified
8965	modified	/features/temperature/propert...		2026-04-23 05:16:25 434254980
8966	modified	/features/temperature/propert...		2026-04-23 05:21:28 399148924
8967	modified	/features/temperature/propert...		2026-04-23 05:21:30 381128104
8968	modified	/features/temperature/propert...		2026-04-23 05:21:32 384150788
8969	modified	/features/temperature/propert...		2026-04-23 05:21:34 387787480
8970	modified	/features/temperature/propert...		2026-04-23 05:21:36 387565338
8971	modified	/features/temperature/propert...		2026-04-23 05:21:38 389997930
8972	modified	/features/temperature/propert...		2026-04-23 05:21:40 391636118
8973	modified	/features/temperature/propert...		2026-04-23 05:21:42 395070223
8974	modified	/features/temperature/propert...		2026-04-23 05:21:44 394542500
8975	modified	/features/temperature/propert...		2026-04-23 05:21:46 397802815
8976	modified	/features/temperature/propert...		2026-04-23 05:23:03 117528430
8977	modified	/features/temperature/propert...		2026-04-23 05:23:05 116058344
8978	modified	/features/temperature/propert...		2026-04-23 05:23:07 117034078
8979	modified	/features/temperature/propert...		2026-04-23 05:23:09 119637134
8980	modified	/features/temperature/propert...		2026-04-23 05:23:11 121482682

Thing Update Detail

Show Only context

```
{
  "revision": 8980,
  "created": "2026-04-23T05:02:59.713091736Z",
  "modified": "2026-04-23T05:23:11.121482682Z",
  "context": {
    "topic": "sensor/temp/things/twz/events/modified",
    "path": "/features/temperature/properties",
    "value": {
      "value": 99.9,
      "timestamp": 1776921791.045335
    }
  }
}
```

Рис. 3. Измененные данные, полученные цифровым двойником

Fig. 3. Modified data received by the digital twin

3.2. REPLAY-АТАКА

Скрипт sniffера подписывается на топик sensors/temperature и сохраняет все сообщения в файл captured_messages.json. Затем скрипт воспроизведения читает файл и публикует сохраненные сообщения в тот же топик с заданной задержкой. Поскольку платформа Ditto не проверяет временные метки, цифровой двойник обновляется устаревшими данными, теряя синхронизацию с реальным объектом (рис. 4). Повторяющиеся временные метки и значения температуры явно видны в интерфейсе Ditto (рис. 5). Уязвимость возникает из-за отсутствия проверки свежести на уровне приложения. TLS не защищает от повторов, поэтому требуется дополнительная логика на стороне двойника.

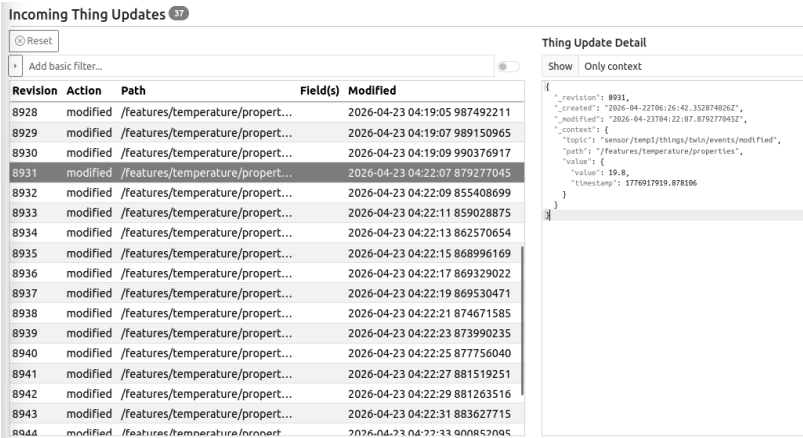


Рис. 4. Данные, полученные цифровым двойником в результате Replay-атаки, первый повтор

Fig. 4. The data obtained by the digital double as a result of the Replay attack, the first replay

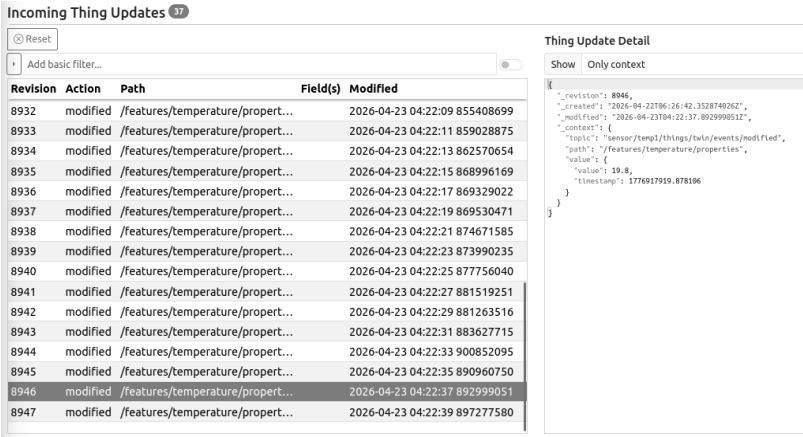


Рис. 5. Данные, полученные цифровым двойником в результате Replay-атаки, второй повтор

Fig. 5. The data obtained by the digital double as a result of the Replay attack, the second replay

3.3. DOS-АТАКА

Многопоточный скрипт (2 потока по 50 сообщений в секунду) публикует фальшивые сообщения в топик sensors/temperature. При атаке на незащищенный канал загрузка CPU контейнера connectivity Eclipse Ditto возрастает до 112 %, контейнера things – до 72 %, легитимные сообщения от реального датчика доставляются с критическими задержками или теряются. Это подтверждает уязвимость брокера и платформы к атакам типа «отказ в обслуживании» даже при сравнительно невысокой интенсивности (100 сообщений в секунду). Рост CPU до 112 % обусловлен парсингом JSON и работой маппера. Атака с интенсивностью 100 сообщений в секунду уже критична. Значит, без фильтрации на сетевом уровне платформа не выдерживает даже умеренного флуда (рис. 6, 7).

```
root@mqtt-server:~/ditto/deployment/docker# docker stats --no-stream
```

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O
PIDS						
abb34114c1ce	docker_nginx_1	0.04%	4.965MiB / 32MiB	15.52%	8.24MB / 8.02MB	53.2kB / 12.3kB
5						
d1c883206faf	docker_things_1	18.16%	484.2MiB / 512MiB	94.57%	162MB / 151MB	53.1MB / 24.3MB
104						
85cb2ac21b0b	docker_connectivity_1	12.18%	565.4MiB / 768MiB	73.62%	171MB / 198MB	15.6MB / 24.1MB
104						
8d0516b0c887	docker_gateway_1	8.48%	459.4MiB / 512MiB	89.72%	169MB / 173MB	51.9MB / 27.8MB
52						
4f7a11ddb10b	docker_things-search_1	6.66%	438.9MiB / 512MiB	85.71%	106MB / 96.4MB	52.5MB / 29.3MB
75						
b0cddf4ac7e7	docker_swagger-ui_1	0.00%	17.45MiB / 32MiB	54.54%	13.2kB / 126B	15.9MB / 1.13MB
5						
d4436cfb8fb0	hivemq	2.34%	709.2MiB / 7.71GiB	8.98%	1.13MB / 1.24MB	2.01MB / 75.5MB
136						
6adae0b349b4	docker_ditto-ui_1	0.00%	4.723MiB / 32MiB	14.76%	21.7kB / 3.03kB	0B / 8.19kB
5						
bed2fada5db7	docker_mongodb_1	1.60%	101.5MiB / 256MiB	39.65%	16.5MB / 31MB	32.8kB / 135MB
107						
3f312abb8a3a	docker_policies_1	5.26%	450.2MiB / 512MiB	87.93%	137MB / 115MB	86MB / 37.7MB
84						

Рис. 6. Проверка нагрузки на сервере с запущенным датчиком

Fig. 6. Checking the load on a server with a running sensor

```
root@mqtt-server:~/ditto/deployment/docker# docker stats --no-stream
```

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O
PIDS						
abb34114c1ce	docker_nginx_1	1.51%	4.965MiB / 32MiB	15.52%	5.66MB / 5.51MB	53.2kB / 12.3kB
5						
d1c883206faf	docker_things_1	112.77%	489.6MiB / 512MiB	95.62%	87.2MB / 88.2MB	53.1MB / 17.9MB
112						
85cb2ac21b0b	docker_connectivity_1	72.54%	585.9MiB / 768MiB	76.28%	91.5MB / 106MB	15.6MB / 19.9MB
112						
8d0516b0c887	docker_gateway_1	23.83%	459.2MiB / 512MiB	89.69%	90.9MB / 90.9MB	51.9MB / 11.9MB
61						
4f7a11ddb10b	docker_things-search_1	16.53%	457MiB / 512MiB	89.26%	57.1MB / 49.8MB	52.5MB / 21.6MB
75						
b0cddf4ac7e7	docker_swagger-ui_1	0.00%	17.45MiB / 32MiB	54.54%	11.9kB / 126B	15.9MB / 1.13MB
5						
d4436cfb8fb0	hivemq	2.92%	711.1MiB / 7.71GiB	9.01%	875kB / 986kB	2.01MB / 74.8MB
140						
6adae0b349b4	docker_ditto-ui_1	0.00%	4.723MiB / 32MiB	14.76%	20.4kB / 3.03kB	0B / 8.19kB
5						
bed2fada5db7	docker_mongodb_1	11.58%	103.6MiB / 256MiB	40.48%	10.7MB / 16.4MB	32.8kB / 76.2MB
107						
3f312abb8a3a	docker_policies_1	5.18%	455.6MiB / 512MiB	88.98%	71.2MB / 59.6MB	86MB / 22.6MB
82						

```
root@mqtt-server:~/ditto/deployment/docker#
```

Рис. 7. Проверка нагрузки на сервере в состоянии идущей DoS-атаки и работающего датчика

Fig. 7. Checking the load on the server in the state of an ongoing DoS attack and a working sensor

4. МЕХАНИЗМЫ ЗАЩИТЫ

4.1. ШИФРОВАНИЕ TLS ДЛЯ MQTT-КАНАЛА

Для защиты канала от перехвата и подмены данных настроен MQTT over TLS на порте 8883. С помощью утилиты keytool сгенерирован самоподписанный сертификат сервера, который импортирован в хранилище ключей NiveMQ. Эмулятор датчика модифицирован для поддержки TLS с проверкой сертификата. При попытке провести MITM-атаку прокси-сервер получает зашифрованный трафик и не может расшифровать содержимое; модификация данных становится невозможной. Захват трафика Wireshark показывает нечитаемый бинарный поток (рис. 8, 9).

No.	Time	Source	Destination	Protocol	Length	Info
1802	0.057430037	192.168.0.155	192.168.0.192	TCP	60	52343 → 1884 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
1803	0.057384937	192.168.0.192	192.168.0.155	TCP	60	1884 → 52343 [SYN, ACK] Seq=0 Ack=1 Win=64240 Len=0 MSS=1460 SACK_PERM=1 WS=128
1804	0.05530992	192.168.0.155	192.168.0.192	TCP	60	52343 → 1884 [ACK] Seq=1 Ack=1 Win=131328 Len=0
1805	0.056491296	192.168.0.155	192.168.0.192	TLSv1.2	216	Client Hello
1806	0.056474896	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=1 Ack=163 Win=64128 Len=0
1814	0.057044812	192.168.0.192	192.168.0.155	TLSv1.2	1350	Server Hello, Certificate, Server Key Exchange, Server Hello Done
1815	0.055376156	192.168.0.155	192.168.0.192	TLSv1.2	147	Client Key Exchange, Change Cipher Spec, Encrypted Handshake Message
1818	0.056432628	192.168.0.192	192.168.0.155	TLSv1.2	1276	New Session Ticket, Change Cipher Spec, Encrypted Handshake Message
1819	0.054182043	192.168.0.155	192.168.0.192	TLSv1.2	113	Application Data
1822	0.054189406	192.168.0.192	192.168.0.155	TLSv1.2	87	Application Data
1826	0.052267511	192.168.0.155	192.168.0.192	TCP	60	52343 → 1884 [ACK] Seq=315 Ack=2554 Win=131328 Len=0
1834	0.054220995	192.168.0.155	192.168.0.192	TLSv1.2	208	Application Data
1837	0.056524996	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=469 Win=64128 Len=0
1839	0.053847926	192.168.0.155	192.168.0.192	TLSv1.2	208	Application Data
1860	0.053936395	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=623 Win=64128 Len=0
1869	0.053771264	192.168.0.155	192.168.0.192	TLSv1.2	208	Application Data
1889	0.053764063	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=777 Win=64128 Len=0
1924	0.053490612	192.168.0.155	192.168.0.192	TLSv1.2	208	Application Data
1925	0.053492411	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=931 Win=64128 Len=0
1955	0.053822263	192.168.0.155	192.168.0.192	TLSv1.2	207	Application Data
1956	0.053758202	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=1084 Win=64128 Len=0
1959	0.053435019	192.168.0.155	192.168.0.192	TLSv1.2	208	Application Data
1987	0.053427509	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=1238 Win=64128 Len=0
1999	0.053192059	192.168.0.155	192.168.0.192	TLSv1.2	208	Application Data
1999	0.053170509	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=1392 Win=64128 Len=0
2016	0.051347476	192.168.0.155	192.168.0.192	TLSv1.2	85	Application Data
2018	0.053192059	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [ACK] Seq=2554 Ack=1424 Win=64128 Len=0
2019	0.051567414	192.168.0.155	192.168.0.192	TCP	60	52343 → 1884 [FIN, ACK] Seq=1423 Ack=2554 Win=131328 Len=0
2020	0.051524139	192.168.0.192	192.168.0.155	TCP	54	1884 → 52343 [FIN, ACK] Seq=2554 Ack=1424 Win=64128 Len=0
2027	0.050431591	192.168.0.155	192.168.0.192	TCP	60	52343 → 1884 [ACK] Seq=1424 Ack=2555 Win=131328 Len=0

Рис. 8. Захват зашифрованного трафика MQTT over TLS

Fig. 8. Capture encrypted MQTT over TLS traffic



Рис. 9. Зашифрованный TLS-трафик

Fig. 9. Encrypted TLS traffic

4.2. ЗАЩИТА ОТ REPLAY-АТАК

В JavaScript-маппере Eclipse Ditto добавлена проверка временной метки входящего сообщения. Если timestamp отличается от текущего времени более чем на 5 секунд, сообщение отбрасывается (функция возвращает null). После внедрения этой проверки все сообщения, воспроизводимые скриптом Replay-атаки, игнорируются. Легитимные сообщения от работающего датчика продолжают обрабатываться штатно. Таким образом, достигается защита от повторного воспроизведения без модификации клиентского ПО.

Incoming JavaScript Mapping

```
1 function mapToDittoProtocolMsg(headers, textPayload, bytePay
2   var json = JSON.parse(textPayload);
3   var currentTime = Math.floor(Date.now() / 1000);
4
5   if (currentTime - json.timestamp > 5) {
6     return null;
7   }
8
9   return Ditto.buildDittoProtocolMsg(
10     'sensor',           // namespace
11     'temp1',            // id
12     'things',
13     'twin',
14     'commands',
15     'modify',
16     '/features/temperature/properties',
17     headers,
18     {
19       value: json.temperature,
20       timestamp: json.timestamp
21     }
22   );
23 }
```

Рис. 10. Модифицированный JavaScript-маппер с проверкой актуальности временной метки

Fig. 10. Modified JavaScript mapper with timestamp validity check

4.3. ЗАЩИТА ОТ DOS-АТАК

На уровне межсетевого экрана (iptables) в цепочке DOCKER-USER применены следующие правила:

- ограничение количества одновременных соединений с одного IP-адреса: `-m connlimit --connlimit-above 1 --connlimit-mask 32 -j DROP`;
- отслеживание подозрительных IP-адресов с помощью модуля recent и блокировка при превышении трех пакетов за 1 секунду;

- ограничение скорости пакетов: `-m hashlimit --hashlimit 3/sec --hashlimit-burst 6 --hashlimit-mode srcip --hashlimit-name dos_limit -j АСCEPT`.

После применения правил атакующий трафик отфильтровывается, а легитимные сообщения датчика проходят без задержек. Загрузка CPU контейнеров возвращается к штатному уровню.

5. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА И ОЦЕНКА ЗАДЕРЖЕК

5.1. ЭФФЕКТИВНОСТЬ ЗАЩИТЫ

Сравнительный анализ сценариев защиты показал, что ни один из механизмов по отдельности не обеспечивает полной защиты. Только их комплексное применение (TLS + проверка временных меток + iptables) снижает риски для всех трех типов атак (табл. 3).

Таблица 3

Table 3

Сравнительный анализ сценариев защиты

Comparative analysis of protection scenarios

Сценарий	MITM-атака успешна	Replay-атака успешна	DoS-атака успешна
Без защиты	Да	Да	Да
Только TLS	Нет для простого MITM в канале	Частично, зависит от модели атакующего	Да
Только защита от Replay	Да	Нет для устаревших сообщений	Да
Только защита от DoS	Да	Да	Снижает успешность
Комплексная защита	Снижает риск	Снижает риск	Снижает риск

5.2. ВЛИЯНИЕ НА ВРЕМЕННЫЕ ХАРАКТЕРИСТИКИ

Для количественной оценки накладных расходов были измерены задержки MQTT publish RTT (Round-Trip Time) в четырех режимах:

- 1) прямое соединение без TLS (базовый режим);

- 2) через MITM-прокси без TLS;
 - 3) прямое соединение с TLS;
 - 4) через MITM-прокси с TLS.
- В каждом режиме 5 серий по 100 измерений (табл. 4).

Таблица 4

Table 4

Статистические характеристики задержек MQTT publish RTT

Statistical characteristics of MQTT delays publish RTT

Сценарий	n	Mean	Median	Std	Min	Max	Q1	Q3	IQR	p90	p95	p99	Выбросы	Выбросы, %	p95/Median
Напрямую без TLS	500	8.53	8.64	3.06	2.60	15.63	5.80	11.04	5.24	12.66	13.21	14.18	0	0.00	1.53
MITM без TLS	500	9.27	9.42	3.04	3.38	17.27	6.65	11.76	5.11	13.51	14.05	14.93	0	0.00	1.49
Напрямую с TLS	500	9.30	9.29	3.12	3.34	21.67	6.64	11.86	5.22	13.28	13.87	15.23	1	0.20	1.49
MITM с TLS	500	10.44	10.65	2.98	4.18	18.48	8.02	12.58	4.57	14.39	15.11	16.54	0	0.00	1.42

Для проверки статистически значимых различий между распределениями задержек MQTT publish RTT во всех четырех экспериментальных сценариях использовался непараметрический критерий Kruskal – Wallis. Для выполнения попарных сравнений между отдельными сценариями применялся критерий Mann – Whitney U с поправкой Holm – Bonferroni на множественные сравнения. Статистический анализ (критерий Краскела – Уоллиса, $p < 0,001$) подтвердил значимые различия между распределениями (табл. 5).

Таблица 5

Table 5

Результаты статистических тестов

Statistical test results

Сравнение	Тест	Статистика	p-value	Вывод
MITM без TLS vs Напрямую без TLS	Mann-Whitney U	$U = 108488.5, Z = -3.62$	< 0.001	Значимо
Напрямую с TLS vs Напрямую без TLS	Mann-Whitney U	$U = 108131.0, Z = -3.69$	< 0.001	Значимо
MITM с TLS vs Напрямую с TLS	Mann-Whitney U	$U = 99348.0, Z = -5.62$	< 0.001	Значимо
MITM с TLS vs MITM без TLS	Mann-Whitney U	$U = 98779.0, Z = -5.74$	< 0.001	Значимо
MITM с TLS vs Напрямую без TLS	Mann-Whitney U	$U = 83585.5, Z = -9.07$	< 0.001	Значимо
Межгрупповое сравнение (все 4 сценария)	Kruskal-Wallis	$H = 86.77, df = 3$	< 0.001	Распределения различны
Сравнение дисперсий	Levene	$W = 1.10, df1 = 3, df2 = 1996$	0.347	Дисперсии равны

Использование TLS увеличивает среднюю задержку на 9 % относительно прямого незащищенного канала, а совместное применение MITM-прокси и TLS – на 23 %. При этом форма распределения сохраняется, доля аномальных выбросов не возрастает, что говорит о стабильности соединения. Полученные задержки остаются приемлемыми для систем мягкого реального времени и не нарушают работоспособность цифрового двойника (рис. 11–13).

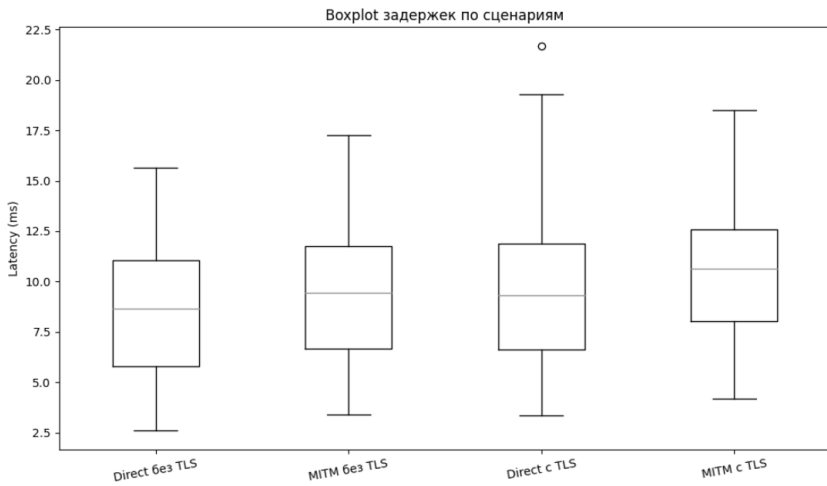


Рис. 11. Boxplot-диаграммы задержек по сценариям

Fig. 11. Boxplot-delay diagrams by scenario

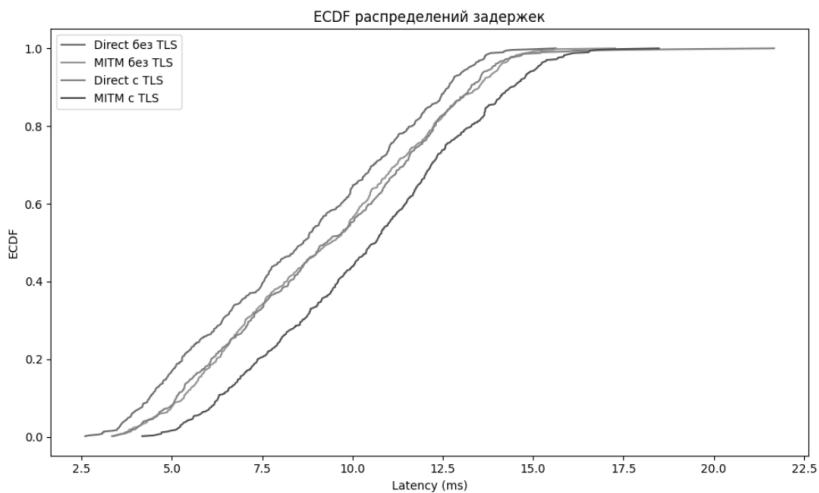


Рис. 12. Распределение задержек ECDF

Fig. 12. ECDF delay Allocation

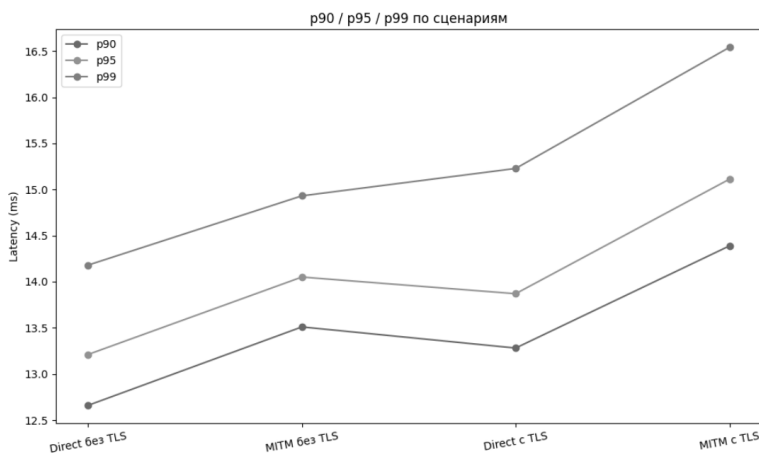


Рис. 13. Перцентили по сценариям

Fig. 13. Percentile scenarios

Данные соглашения показывают стабильную динамику показателей.

ЗАКЛЮЧЕНИЕ

В работе исследованы уязвимости каналов связи цифровых двойников в распределенных вычислительных системах на примере передачи телеметрии по протоколу MQTT. Экспериментально подтверждено, что незащищенный MQTT-канал (порт 1883) подвержен трем типам атак: MITM с подменой данных, Replay-атаке и DoS-атаке интенсивностью всего 100 сообщений в секунду, которая приводит к перегрузке брокера и рассинхронизации цифрового двойника. Предложен и апробирован комплекс практических мер защиты: шифрование MQTT over TLS (порт 8883) с проверкой сертификата сервера, что снижает вероятность чтения и модификации трафика при помощи MITM-прокси; модификация JavaScript-маппера Eclipse Ditto с отбрасыванием сообщений старше пяти секунд, значительно снижающая вероятность Replay-атаки; настройка правил iptables в цепочке DOCKER-USER (модули connlimit, hashlimit, recent), фильтрующая высокочастотный флуд без влияния на легитимный датчик.

Статистический анализ задержек MQTT publish RTT показал, что использование TLS увеличивает среднее время отклика на 9 %, а совместное применение TLS и промежуточного прокси – на 23 % относительно незащищенного канала, при этом форма распределения задержек сохраняется, что допустимо

для систем мягкого реального времени. Таким образом, комплексное применение TLS, проверки временных меток и ограничений на уровне межсетевого экрана обеспечивает надежную защиту каналов связи цифровых двойников от основных сетевых атак без нарушения их работоспособности.

СПИСОК ЛИТЕРАТУРЫ

1. A survey on digital twin for industrial internet of things: Applications, technologies and tools / H. Xu, J. Wu, Q. Pan, X. Guan, M. Guizani // *IEEE Communications Surveys & Tutorials*. – 2023. – Vol. 25 (4). – P. 2569–2598.
2. Attaran S., Attaran M., Celik B.G. Digital Twins and Industrial Internet of Things: Uncovering operational intelligence in industry 4.0 // *Decision Analytics Journal*. – 2024. – Vol. 10. – P. 100398. – DOI: 10.1016/j.dajour.2024.100398.
3. Soudan B. Cybersecurity of digital twins in industrial IoT environments // 2024 Advances in Science and Engineering Technology International Conferences (ASET). – IEEE, 2024. – P. 1–6. – DOI: 10.1109/ASET60340.2024.10708640.
4. MQTT vulnerabilities, attack vectors and solutions in the Internet of Things (IoT) / A.J. Hintaw, S. Manickam, M.F. Aboalmaaly, S. Karuppayah // *IETE Journal of Research*. – 2023. – Vol. 69 (6). – P. 3368–3397. – DOI: 10.1080/03772063.2021.1912651.
5. Dinculeană D., Cheng X. Vulnerabilities and limitations of MQTT protocol used between IoT devices // *Applied Sciences*. – 2019. – Vol. 9 (5). – P. 848. – DOI: 10.3390/app9050848.
6. Wong H.C. Man-in-the-Middle Attacks on MQTT based IoT networks. – Missouri University of Science and Technology, 2022.
7. Empl P., Pernul G. Digital-twin-based security analytics for the internet of things // *Information*. – 2023. – Vol. 14 (2). – P. 95. – DOI: 10.3390/info14020095.
8. Kampourakis K.E., Gkioulos V., Katsikas S. Cybersecurity Digital Twins for industrial systems: from literature synthesis to framework design // *Information*. – 2026. – Vol. 17 (3). – P. 286. – DOI: 10.3390/info17030286.
9. Методика оценки угроз безопасности информации: утв. ФСТЭК России 05.02.2021. – М., 2021.
10. Банк данных угроз безопасности информации ФСТЭК России. – URL: <https://bdu.fstec.ru/threat> (дата обращения: 05.20.2026).

Сильченко Никита Сергеевич, лаборант кафедры защиты информации Новосибирского государственного технического университета. Область научных интересов – безопасность каналов связи цифровых двойников, противодействие сетевым атакам. E-mail: silchenko.2022@stud.nstu.ru

Пермяков Руслан Анатольевич, старший преподаватель кафедры защиты информации Новосибирского государственного технического университета. Основное направление научных исследований – обеспечение информационной безопасности распределенных вычислительных систем и цифровых двойников. E-mail: permyakov@corp.nstu.ru

DOI: 10.17212/2782-2230-2026-3-24-40

Protection of digital twin communication channels from attacks in distributed computing systems*

N.S. Silchenko¹, R.A. Permyakov²

¹ *Novosibirsk State Technical University, 136 Nemirovich-Danchenko Street, Novosibirsk, 630087, Russian Federation, Laboratory assistant at the Department of Information Security. E-mail: silchenko.2022@stud.nstu.ru*

² *Novosibirsk State Technical University, 136 Nemirovich-Danchenko Street, Novosibirsk, 630087, Russian Federation, Senior Lecturer at the Department of Information Security. E-mail: permyakov@corp.nstu.ru*

Digital twin technology is being actively introduced into the industrial Internet of Things to monitor and manage complex cyber-physical systems in real time. The two-way nature of data exchange between a physical object and its virtual copy generates additional attack vectors, with attacks on communication channels using the lightweight MQTT protocol being particularly dangerous. The paper develops a model of information security threats for digital twin communication channels in accordance with the FSTEC methodology of Russia, and identifies six relevant threats, including data substitution, replay, and denial of service. An experimental stand has been created based on a temperature sensor emulator, the HiveMQ MQTT broker, and the Eclipse Ditto digital twin platform deployed in Docker containers. MITM attack scenarios with data substitution, Replay attacks and DoS attacks have been implemented. It is shown that an unsecured MQTT channel is vulnerable to all listed threats. A set of protection mechanisms is proposed: TLS channel encryption, timestamp verification in a JavaScript mapper, and iptables rules to limit traffic intensity. The effectiveness of the proposed measures has been experimentally confirmed: a MITM attack becomes impossible when the certificate is checked correctly, a Replay attack is prevented for the scenario under consideration by discarding outdated messages, and a DoS attack is significantly reduced due to filtering at the firewall level.

Keywords: digital twin, MQTT, TLS, MITM, Replay attack, DoS attack, iptables, Eclipse Ditto, HiveMQ, distributed computing systems, threat model

* Received 02 July 2026.

REFERENCES

1. Xu H., Wu J., Pan Q., Guan X., Guizani M. A survey on digital twin for industrial internet of things: Applications, technologies and tools. *IEEE Communications Surveys & Tutorials*, 2023, vol. 25 (4), pp. 2569–2598.
2. Attaran S., Attaran M., Celik B.G. Digital Twins and Industrial Internet of Things: Uncovering operational intelligence in industry 4.0. *Decision Analytics Journal*, 2024, vol. 10, p. 100398. DOI: 10.1016/j.dajour.2024.100398.
3. Soudan B. Cybersecurity of digital twins in industrial IoT environments. *2024 Advances in Science and Engineering Technology International Conferences (ASET)*. IEEE, 2024, pp. 1–6. DOI: 10.1109/ASET60340.2024.10708640.
4. Hintaw A.J., Manickam S., Aboalmaaly M.F., Karuppayah S. MQTT vulnerabilities, attack vectors and solutions in the Internet of Things (IoT). *IETE Journal of Research*, 2023, vol. 69 (6), pp. 3368–3397. DOI: 10.1080/03772063.2021.1912651.
5. Dinculeană D., Cheng X. Vulnerabilities and limitations of MQTT protocol used between IoT devices. *Applied Sciences*, 2019, vol. 9 (5), p. 848. DOI: 10.3390/app9050848.
6. Wong H.C. *Man-in-the-Middle Attacks on MQTT based IoT networks*. Missouri University of Science and Technology, 2022.
7. Empl P., Pernul G. Digital-twin-based security analytics for the internet of things. *Information*, 2023, vol. 14 (2), p. 95. DOI: 10.3390/info14020095.
8. Kampurakis K.E., Gkioulos V., Katsikas S. Cybersecurity Digital Twins for industrial systems: From literature synthesis to framework design. *Information*, 2026, vol. 17 (3), p. 286. DOI: 10.3390/info17030286.
9. FSTEC of Russia. *Metodika otsenki ugroz bezopasnosti informatsii* [Methodology for assessing information security threats]. Moscow, 2021.
10. FSTEC of Russia. *Information security threats database*. (In Russian). Available at: <https://bdu.fstec.ru/threat> (accessed 20.05.2026).

Для цитирования:

Сильченко Н.С., Пермяков Р.А. Защита каналов связи цифровых двойников от атак в распределенных вычислительных системах // Безопасность цифровых технологий. – 2026. – № 3 (122). – С. 24–40. – DOI: 10.17212/2782-2230-2026-3-24-40.

For citation:

Silchenko N.S., Permyakov R.A. Zashchita kanalov svyazi tsifrovyykh dvoynikov ot atak v raspredelennykh vychislitel'nykh sistemakh [Protection of digital twin communication channels from attacks in distributed computing systems]. *Bezopasnost' tsifrovyykh tekhnologii = Digital Technology Security*, 2026, no. 3 (122), pp. 24–40. DOI: 10.17212/2782-2230-2026-3-24-40.

*МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ*

УДК 004

DOI: 10.17212/2782-2230-2026-3-41-63

**РЕАЛИЗАЦИЯ КРИПТОАНАЛИЗА ШИФРА
ПРОСТОЙ ЗАМЕНЫ С ИСПОЛЬЗОВАНИЕМ
ЗНАКОВЫХ ТРИГРАММ***

Ю.А. КОТОВ¹, А.П. БАШКИРОВ²

¹ РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, кандидат физико-математических наук, доцент, доцент кафедры защиты информации. E-mail: kotov@corp.nstu.ru

² РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, лаборант кафедры защиты информации. E-mail: kaf_zi@corp.nstu.ru

В работе рассматривается адаптация метода Якобсена для криптоанализа шифра простой замены с использованием знаковых триграмм применительно к русскоязычным текстам. Проведена оценка устойчивости эталона: стабильности частот и рангов наиболее частотных триграмм, полноты покрытия множества ненулевых триграмм и геометрического сходства их распределения по пластинам. Для сопоставимости частот применена нормировка, экспериментально определена оптимальная стратегия нормировки в зависимости от объема анализируемого текста. Метод реализован программно и протестирован на независимой выборке научно-популярных и художественных текстов. Сравнительный анализ с биграммной реализацией показал, что триграммный подход обеспечивает существенно меньшую ошибку идентификации (вплоть до нуля при объемах свыше 6000 символов) при приемлемых вычислительных затратах. Полученные результаты позволяют количественно оценить минимальный объем шифротекста, при котором атака на основе триграмм становится результативной, и подтверждают практическую ценность триграммного эталона для статистического криптоанализа русскоязычных текстов.

Ключевые слова: триграммы, частотный анализ, шифр простой замены, метод Якобсена, эталон частот, нормировка

* Статья получена 14 июля 2026 г.

ВВЕДЕНИЕ

Частотный анализ текстов, основанный на статистических закономерностях естественного языка, является одним из инструментов криптоанализа и реализуется через сравнение частот символов и их сочетаний – униграмм, биграмм и триграмм – с эталонными характеристиками языка [1]. Наиболее наглядно его применение демонстрируется на модели криптоанализа шифра простой замены, в котором все статистические свойства открытого текста сохраняются в шифротексте [2]. Вместе с тем статистические характеристики буквенных последовательностей, в том числе триграммы, находят применение и в смежных задачах – в частности, в задаче установления принадлежности анонимного текста конкретному автору и других задачах.

Переход от анализа отдельных символов к биграммам и триграммам существенно повышает точность криптоанализа за счет учета контекстуальных связей языка [3]. При дальнейшем увеличении длины сочетаний знаков можно перейти к использованию словарного метода, который дает меньшую ошибку [4]. Но частотные методы обладают рядом практических преимуществ перед словарными методами анализа текстов. Во-первых, из текста могут быть удалены пробелы и знаки препинания (что характерно, например, при шифровании), что делает словарный подход в таких случаях неприменимым без предварительного восстановления разбивки текста на слова. Во-вторых, частотный метод не зависит от конкретного языка: для его применения достаточно располагать эталонным распределением сочетаний знаков целевого языка.

В работе [1] использование триграмм в метаэвристических алгоритмах криптоанализа признается наиболее эффективным подходом.

С другой стороны, некоторые авторы отмечают, что при работе с русскоязычными текстами возникают дополнительные сложности, обусловленные морфологическим богатством русского языка: разнообразие словоформ ведет к большему варьированию статистики буквенных сочетаний [5]. Это может повышать требования к объему шифротекста, необходимому для получения репрезентативной триграммной статистики, поэтому зависимость успешности триграммного криптоанализа от длины шифротекста требует отдельного экспериментального исследования [5].

Применение триграммного анализа требует наличия эталона частот триграмм – набора эталонных частот встречаемости трехбуквенных последовательностей, характерных для данного языка. Для русскоязычных текстов подобный эталон биграмм был применен в методе Якобсена [2], однако аналогичного эталона триграмм в доступной литературе обнаружено не было. Настоящая работа направлена на восполнение этого пробела и сравнитель-

ную оценку эффективности триграммного подхода по отношению к биграммному.

Целью данной работы является адаптация метода Якобсена при криптоанализе шифра простой замены с использованием знаковых триграмм для русскоязычных текстов.

1. ФОРМИРОВАНИЕ ЭТАЛОНА ТРИГРАММ

1.1. ХАРАКТЕРИСТИКА ИСХОДНОГО ТЕКСТА И ПРОЦЕДУРА ПРЕДОБРАБОТКИ

Для криптоанализа шифра замены необходим эталон частот триграмм, отражающий вероятности их встречаемости в естественном языке. Готового эталона в литературе не обнаружено, поэтому был создан собственный.

В качестве корпуса использовалась «выборка 2» из работы [2]. Этот корпус уже применялся ранее, что обеспечивает корректное сравнение с результатами биграммных методов. Выборка включает 100 текстов вузовских учебных пособий разных авторов из различных областей (математика, химия, физика, машиностроение и др.). Из каждого текста случайным образом выделялись последовательные фрагменты различной длины, начинающиеся со случайного знака. Такой подход моделирует «произвольный» текст.

Русский алфавит приведен к 31 букве: мягкий и твердый знаки объединены в «Ь», «Е» и «Ё» – в «Е». Из текстов исключены пробелы, знаки препинания, цифры и прочие небуквенные символы, в статистике участвуют только буквы.

Эталон триграмм формировался суммированием всех триграмм по корпусу только для фрагментов с полным алфавитом. Количественные характеристики обработанных групп текстов приведены в табл. 1.

Таблица 1

Table 1

Характеристика групп текстов

Characteristics of groups of texts

G	D	S	K	V
1	200...2000	200	548	669 583
2	2000...10 000	2000	481	2 554 927
3	10 000...110 000	20 000	275	14 336 924
4	100 000...350 000	50 000	48	9 411 829
Итого:			1352	26 973 263

В табл. 1 столбец G содержит номер группы текстов, столбец D – диапазон объемов текстов в символах, столбец S – шаг изменения объема текстов в данном диапазоне в символах, столбец K – количество текстов, столбец V – общий объем символов текстовой группы.

Важным вопросом при формировании эталона триграмм является вопрос его устойчивости как характеристики языка. Для этого следует оценить стабильность частот триграмм, полноту покрытия триграмм, стабильность «рисунка» триграмм, т. е. сходства расположения триграмм на соответствующих срезах трехмерного массива частот встречаемости триграмм при переходе от одной группы текстов к другой, а также решить вопрос нормировки эталона.

1.2. АНАЛИЗ СТАБИЛЬНОСТИ ЧАСТОТ ТРИГРАММ

Рассмотрим задачу о стабильности частот триграмм для русскоязычных текстов.

На рис. 1 представлен график распределения частот встречаемости триграмм для суммарной статистики по всем группам текстов, где для каждого значения частоты показано, сколько именно триграмм встретилось с данной частотой.

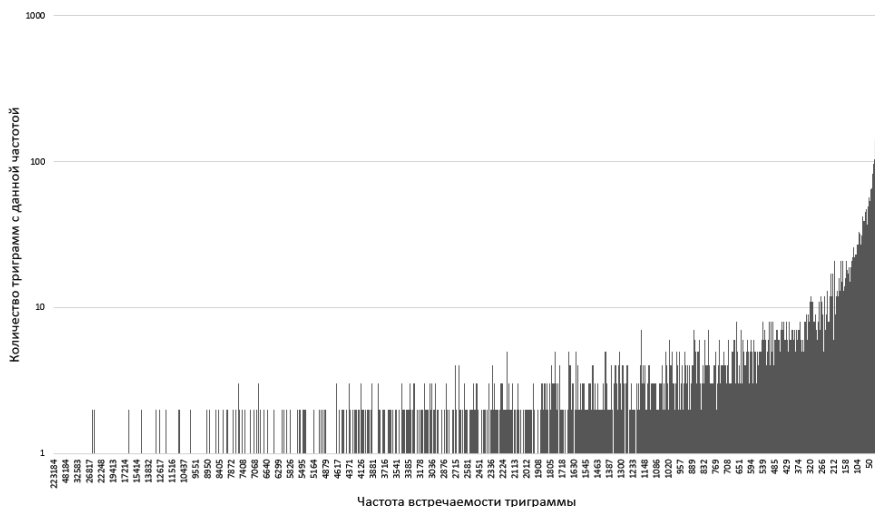


Рис. 1. Распределение частот встречаемости триграмм

Fig. 1. Frequency distribution of trigrams

Из рис. 1 видно, что в области высоких частот каждой частоте, как правило, соответствует единственная триграмма. При переходе к низким частотам одна и та же частота может относиться уже к сотням различных триграмм. Для проверки стабильности частотных характеристик рассмотрено 20 наиболее часто встречающихся триграмм по всем группам текстов. Такой выбор позволяет оперировать триграммами без частотных коллизий и исключить редкие сочетания, часто являющиеся случайными.

Далее под суммарной статистикой понимается частотная статистика триграмм, вычисленная по всему корпусу выборки 2 (без разделения на группы) объемом 26 973 263 символа.

В табл. 2 приведены 20 наиболее часто встречаемых триграмм и их частоты для групп текстов 2, 3, 4, а также для суммарной статистики.

Таблица 2

Table 2

Статистика по наиболее часто встречаемым триграммам**Statistics on the most common trigrams**

Триграмма	Суммарная статистика	Группа текстов 2	Группа текстов 3	Группа текстов 4
ЕНИ	223 184	21 640	118 048	77 984
ОСТ	169 526	15 244	90 139	60 058
НИЯ	129 043	12 155	68 613	44 952
СТВ	103 505	8801	53 200	39 055
НИЕ	99 309	9521	51 551	35 856
ПРИ	96 515	9053	51 446	33 619
СТИ	93 782	7343	48 372	36 090
ПРО	92 489	8700	50 296	31 260
ОГО	91 874	9077	48 664	31 750
НОС	90 532	8374	47 583	32 444
ЛЬН	89 514	9302	50 040	27 702
ТСЯ	85 629	7944	46 236	29 251
АНИ	79 783	7295	41 526	28 981
ЕТС	76 305	6803	40 654	26 896
ЛЕН	74 373	6615	39 389	26 800
ИЧЕ	73 415	6473	38 312	26 850
ЕНН	73 128	6539	35 499	29 274
РЕД	73 108	7037	39 330	25 050
ПОЛ	72 363	7567	39 335	23 514
ЧЕС	71 079	6155	36 582	26 626

Из табл. 2 видно, что во всех группах текстов состав топ-20 триграмм практически идентичен: триграммы ЕНИ, ОСТ, НИЯ, СТВ, НИЕ, ПРИ и другие неизменно входят в число наиболее встречаемых.

Однако важно оценить не только состав, но и порядок следования триграмм по частоте. Для этого вычислены разности рангов каждой триграммы в группе текстов и ранга в суммарной статистике. Положительная разность означает, что в группе текстов триграмма опустилась ниже, отрицательная – поднялась выше. Полученные значения (для групп текстов 1–4) визуализированы на рис. 2 в виде тепловой карты [6].

Триграмма	Группа текстов 1	Группа текстов 2	Группа текстов 3	Группа текстов 4
ЕНИ	0	0	0	0
ОСТ	0	0	0	0
НИЯ	0	0	0	0
СТВ	1	4	0	0
НИЕ	3	-1	0	1
ПРИ	0	1	0	1
СТИ	6	6	3	-2
ПРО	1	1	-1	2
ОГО	-2	-3	0	0
НОС	1	0	1	-2
ЛЬН	-7	-6	-3	4
ТСЯ	-2	-1	0	0
АНИ	-1	1	0	0
ЕТС	0	4	0	3
ЛЕН	11	4	0	4
ИЧЕ	3	5	2	2
ЕНН	1	3	5	-6
РЕД	3	-1	-1	3
ПОЛ	-4	-7	-3	3
ЧЕС	0	5	-1	0

Рис. 2. Разность рангов триграмм

Fig. 2. The difference in the ranks of trigrams

Анализ тепловой карты (рис. 2) показывает, что разности рангов в группах 3 и 4 не превышают трех-четырёх позиций для большинства триграмм, а для многих равны нулю. В группе 2 разности выше (до семи), в группе 1 достигают 11, что отражает более сильные отклонения в коротких фрагментах. Тем не менее состав топ-20 триграмм практически неизменен во всех группах.

Если бы частоты определялись преимущественно случайными факторами, набор наиболее частотных сочетаний менялся бы хаотически. Регулярное воспроизведение одного и того же ядра высокочастотных триграмм свидетельствует о наличии устойчивых языковых закономерностей, пригодных для построения эталона.

Триграммы образуют трехмерное распределение. Двумерный срез такого распределения в той или иной плоскости назовем «пластиной».

Рассмотрены также наиболее частотные триграммы в пластинах по строкам: при фиксированной третьей букве в каждой строке (при фиксированной первой букве) находилась максимальная по частоте триграмма. Результат аналогичен предыдущему: триграммные максимумы схожи во всех группах, что подтверждает устойчивость триграммного распределения.

2. ПОЛНОТА ПОКРЫТИЯ ТРИГРАММ

Для уточнения характера распределения триграмм в русскоязычных текстах рассмотрим полноту их представления. Важным параметром статистической структуры языка является количество реально встречающихся триграмм – комбинаций из трех букв, появляющихся в тексте хотя бы один раз. Теоретически возможно 29 791 различных триграмм, однако реальный язык использует лишь их часть. Сравнение групп текстов разного объема позволяет оценить зависимость полноты покрытия от количества проанализированного материала.

Рассмотрено распределение количества ненулевых триграмм для каждой из 31 пластины. На рис. 3 представлен график количества ненулевых триграмм по пластинам для всех групп текстов. Для наглядности буквы на оси абсцисс упорядочены не по частоте встречаемости, а по числу ненулевых триграмм, в которых они занимают третье место.

На рис. 3 и далее цифрами 1–4 обозначены группы текстов (см. табл. 1), цифрой 5 – суммарная статистика. Графики показывают, что для каждой буквы число ненулевых триграмм стабильно растет от группы 1 к группе 3: чем больше текста, тем больше редких сочетаний проявляется. Исключение – группа 4, где значения ниже, чем в группе 3, хотя фрагменты длиннее, но текстов меньше (48 против 275). Это свидетельствует о том, что суммарное разнообразие триграмм при малом числе источников может быть ниже, чем при большом количестве коротких отрывков. Суммарная статистика превышает показатели любой группы, аккумулируя все триграммы.

В табл. 3 приведено количество ненулевых триграмм для каждой группы и процент от теоретического максимума (29 791).

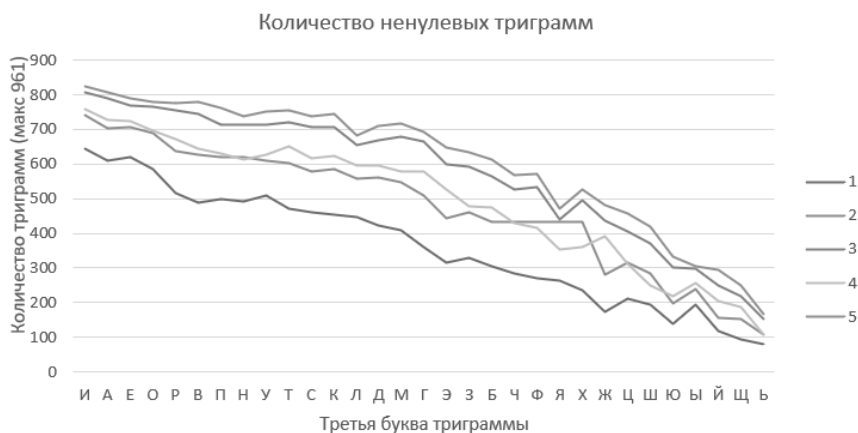


Рис. 3. Количество ненулевых триграмм по пластинам

Fig. 3. The number of non-zero trigrams on the plates

Таблица 3

Table 3

Количество ненулевых триграмм по группам текстов

Number of non-zero trigrams by text groups

Номер группы текстов	Количество ненулевых триграмм	Процент от теоретического максимума
1	11 205	37,61 %
2	14 449	48,50 %
3	17 765	59,63 %
4	15 303	51,37 %
Суммарная статистика	18 796	63,09 %

Если принять суммарную статистику за 100 %, то группа текстов 1 покрывает 59,6 % эталонного разнообразия, группа текстов 2 – 76,9 %, группа текстов 3 – 94,5 %, группа текстов 4 – 81,4 %. Таким образом, уже группа

текстов 3 обеспечивает практически полное покрытие от того, что дает суммарный эталон. Дальнейшее увеличение объема за счет добавления групп текстов 1, 2 и 4 дает прирост лишь на 5,5 %. Это свидетельствует о достижении насыщения: большинство реально существующих в языке триграмм является уже при анализе текстов группы 3.

3. ПРОВЕРКА СТАБИЛЬНОСТИ ГЕОМЕТРИЧЕСКОГО РАСПРЕДЕЛЕНИЯ

Если эталон является устойчивым «усредненным портретом» языка, то его множество ненулевых триграмм должно быть близко к множествам, полученным из больших независимых выборок (группы 3 и 4). Для количественной оценки этой близости применяется бинарное сравнение, при котором учитывается только факт наличия триграммы, но не ее частота.

Для каждой группы текстов (G) и для каждой пластины (c) была поставлена в соответствие бинарная матрица $M_G^{(c)}$, где, как и в пластинах, строки соответствуют первой букве триграммы, столбцы – второй. Элемент матрицы определяется следующим образом:

$$M_G^{(1)}(i, j) = \begin{cases} 1, & \text{если триграмма } (i, j, c) \text{ встретилась хотя бы раз,} \\ 0 & \text{в противном случае,} \end{cases} \quad (1)$$

где i, j – индексы букв.

Матрица разности для двух групп текстов A и B определяется по формуле

$$D_{A,B}^{(c)} = M_A^{(c)} - M_B^{(c)}. \quad (2)$$

Элементы $D_{A,B}^{(c)}(i, j)$ могут принимать три значения:

- 1) 0 – триграмма (i, j, c) присутствует либо отсутствует в обеих группах текстов;
- 2) 1 – триграмма есть только в группе текстов A ;
- 3) -1 – триграмма есть только в группе текстов B .

Далее для каждой пластины подсчитывается количество уникальных триграмм по формулам:

$$S_1^{(c)} = \sum \left(D_{A,B}^{(c)}(i, j) = 1 \right), \quad S_{-1}^{(c)} = \sum \left(D_{A,B}^{(c)}(i, j) = -1 \right), \quad (3)$$

где $S_1^{(c)}$ – количество триграмм, уникальных для группы текстов A ; $S_{-1}^{(c)}$ – количество триграмм, уникальных для группы текстов B .

Абсолютные величины S_{-1} и S_1 зависят от общего числа ненулевых триграмм в каждой группе. Чтобы получить характеристику, позволяющую сравнивать различные объемы текстов, вычисляется значение по формуле

$$R_{A,B}^{(c)} = \frac{S_1^{(c)} + |S_{-1}^{(c)}|}{N_A^{(c)} + N_B^{(c)}}, \quad (4)$$

где $N_A^{(c)} = \sum \left(M_A^{(c)}(i, j) \right)$ – общее количество ненулевых триграмм в пластине c для группы текстов A .

Если $R = 0$, множества ненулевых триграмм полностью совпадают; если $R = 1$ – множества не пересекаются. Промежуточные значения отражают долю уникальных триграмм в объединенном множестве.

Рассматривались попарные сравнения независимых групп текстов и сравнение каждой группы с суммарной статистикой. Попарное сравнение выполнялось с группой 3, обладающей наибольшим объемом (более 14 млн символов). На рис. 4 представлены графики значений R для различных сочетаний групп текстов; по оси абсцисс отложены буквы пластин.

На графике, представленном на рис. 4, видно, что для всех пластин значения R уменьшаются от пары (3 + 1) к паре (3 + 4). Например, для буквы О: $0,1376 \rightarrow 0,0646 \rightarrow 0,0562$; для буквы Ж: $0,4525 \rightarrow 0,2580 \rightarrow 0,1325$. Это указывает на то, что с ростом объема текста множество ненулевых триграмм стремится стать идентичным. Для самых часто встречаемых букв даже в группе текстов 1 значения R составляют $0,12 \dots 0,13$, т. е. $87 \dots 88\%$ триграмм совпадают с группой текстов 3. Для редких букв R может достигать $0,3 \dots 0,4$ даже в группе текстов 4, что говорит о высокой вариативности сочетаемости редких букв.

Для оценки того, за счет какой группы текстов происходит расхождение, на рис. 5 показан вклад каждой группы текстов (1 и 3) отдельно для сравнения их друг с другом.



Рис. 4. Результат сравнения для ненормированных данных

Fig. 4. Comparison result for non-normalized data



Рис. 5. Сравнение вклада групп текстов по отдельности

Fig. 5. Comparing the contribution of groups of texts individually

Из рис. 5 видно, что триграмм, присутствующих только в группе текстов 3, на порядок больше. Это говорит об отсутствии многих реально существующих сочетаний в малых выборках небольших фрагментов текстов.

Дополнительно было проведено сравнение каждой группы текстов с суммарной статистикой. Результаты для ненормированных данных приведены на рис. 6.



Рис. 6. Сравнение с суммарной статистикой

Fig. 6. Comparison with summary statistics

Сравнение с суммарной статистикой дает оптимальную оценку сходства: любая триграмма группы входит в суммарную статистику. Поэтому R здесь – доля триграмм эталона, отсутствующих в данной группе. Для группы 1 она составляет в среднем 13...45 % (в зависимости от пластины), для группы 3 – менее 1...2 %.

Для каждой группы вычислена разность R при сравнении с суммарной статистикой и с группой 3; разности просуммированы по всем пластинам. Для групп 2 и 4 получено уменьшение R , что указывает на большую близость суммарной статистики к этим группам, чем у самой большой независимой группы. Для группы 1, напротив, значение R увеличилось из-за малого объема фрагментов. В целом результаты подтверждают, что суммарная статистика наиболее полно отражает языковые закономерности и является оправданным выбором в качестве эталона.

3.1. НОРМИРОВКА ДАННЫХ

Корпус текстов для построения эталона имеет объем более 26 млн символов. При таком объеме в эталоне появляются очень редкие триграммы – результат случайных сочетаний, специфических терминов или опечаток. Половина всех триграмм в суммарной статистике имеет частоту встречаемости более 12 735. Большое число редких триграмм затрудняет оценку различия при сопоставлении эталона с анализируемым текстом ограниченного объема, поскольку они практически отсутствуют в последнем.

Для приведения частот к сопоставимому виду применена нормировка к объему 5000 символов по формуле

$$f_{\text{норм}} \approx \frac{x}{N} 5000, \quad (5)$$

где x – количество вхождений конкретной триграммы в анализируемом корпусе; N – общее число символов (букв) в этом корпусе.

Полученное значение округляется до ближайшего целого.

При нормировке сохраняется порядок: если частота одной триграммы была выше другой, после нормировки она не станет меньше, но может сравняться. Нормировка статистики по всем текстам выполнялась после суммирования данных по группам.

После нормировки и округления возникает вопрос: не теряется ли важная информация при отбрасывании или округлении редких триграмм? Для ответа необходимо проверить, насколько сильно различается распределение триграмм в разных группах текстов и как влияет нормировка с округлением на степень их сходства.

С использованием нормированных частот для каждой группы рассчитан коэффициент расхождения R по формуле (4). На рис. 7 приведены распределения ненулевых триграмм для нормированных и ненормированных данных.

Как видно из графиков на рис. 7, нормировка с округлением занижает коэффициент расхождения R , причем сильнее всего для пар групп с изначально большими различиями. Например, при сравнении группы 2 и группы 3 нормировка уменьшает R гораздо заметнее, чем при сравнении группы 3 и группы 4.

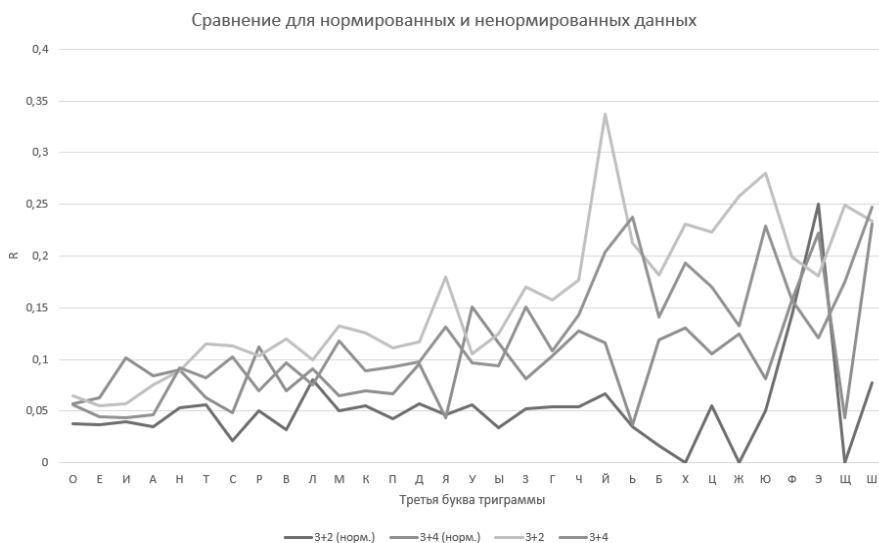


Рис. 7. Сравнение нормированных и ненормированных данных

Fig. 7. Comparison of normalized and non-normalized data

Окончательно влияние нормировки рассмотрим в разделе 4.

4. РЕАЛИЗАЦИЯ МЕТОДА ЯКОБСЕНА КРИПТОАНАЛИЗА ШИФРА ПРОСТОЙ ЗАМЕНЫ НА ОСНОВЕ ЗНАКОВЫХ ТРИГРАММ

4.1. АДАПТАЦИЯ МЕТОДА ЯКОБСЕНА ДЛЯ ТРИГРАММ

Формально N -граммой на алфавите A , состоящем из конечного числа символов, называется строка из N символов, принадлежащих алфавиту A [7].

Модель триграмм базируется на наборах:

$$G_i = \{g_{i1}, g_{i2}, g_{i3}, \theta(g_{i1}, g_{i2}, g_{i3})\}, \quad (6)$$

где $g_{i1} \in A_g$, $g_{i2} \in A_g$, $g_{i3} \in A_g$ – символы из словарного алфавита A_g , (g_{i1}, g_{i2}, g_{i3}) – частота встречаемости триграммы, т. е. количество повторений трехсимвольной цепочки (g_{i1}, g_{i2}, g_{i3}) . В анализируемом тексте частоты

триграмм обычно называют частотными характеристиками, поскольку их легко привести к вероятностной форме нормированием к общему числу символов или на число определенных символов. На практике чаще используют абсолютные частоты (число повторов) [8].

При идентификации знаков текста на основе триграммного распределения частот применяется обобщение метода Якобсена, изначально предложенного для биграмм [2]. Алгоритм основан на минимизации целевой функции вида

$$W(T) = \left(\sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N |t_{ijk} - b_{ijk}| \right) - \delta, \quad (7)$$

где t_{ijk} и b_{ijk} – значения частот триграмм анализируемого текста и эталона соответственно; δ – поправка для устранения двойного учета триграмм; N – размер алфавита.

Поиск минимального значения $W(T)$ выполняется процедурой упорядочивания знаков с переменным шагом. Начальный порядок алфавита задается сортировкой по убыванию частот отдельных знаков в анализируемом тексте, эталонный порядок – аналогичной сортировкой частот в эталоне. Далее последовательно просматриваются пары знаков: сначала соседние, затем отстоящие на один, два знака и более, вплоть до пары из первого и последнего (максимальный шаг $N - 1$). Множество пар с одинаковым расстоянием между индексами образует уровень просмотра, а индекс начального элемента пары – ее место. Просмотр на каждом уровне начинается с первого символа.

Для каждой пары вычисляются W_0 при текущем порядке и W_1 после условной перестановки этих знаков. Перестановка влечет одновременную перестановку соответствующих слоев во всех трех измерениях куба частот. Если выполняется условие

$$W_0 > W_1, \quad (8)$$

то пара переставляется окончательно, и просмотр возвращается к нулевому уровню (шаг = 1) с самого начала. В противном случае знаки возвращаются на исходные позиции, и просмотр продолжается в прежнем порядке.

Процесс повторяется, пока при полном проходе по всем уровням не будет выполнено ни одной перестановки. Таким образом, метод Якобсена для триграмм представляет собой сортировку с переменным шагом, адаптированную к минимизации трехмерного распределения частот, и является естественным обобщением двумерного варианта.

5. СРАВНЕНИЕ БИГРАММНОЙ И ТРИГРАММНОЙ РЕАЛИЗАЦИИ МЕТОДА ЯКОБСЕНА ДЛЯ РУССКОЯЗЫЧНЫХ ТЕКСТОВ

Сформированный по выборке 2 эталон триграмм требует экспериментальной проверки на реальных текстах. Для этого использована «выборка 1» из работы [2], тексты которой не участвовали в формировании эталона. Выборка 1 отличается составом и способом формирования фрагментов, что позволяет оценить универсальность эталона для языка.

Выборка 1 составлена из научно-популярных и художественных текстов различных жанров и авторов. В отличие от выборки 2 (случайные фрагменты), здесь применен принцип вложенности: после удаления пробелов выделялись фрагменты большей длины, из которых затем формировались меньшие фрагменты путем последовательного удаления постоянного объема знаков. Начала всех фрагментов одного текста совпадали. Это позволяет оценить зависимость погрешности метода от объема текста.

Объемы фрагментов распределены по четырем группам, аналогичным группам выборки 2:

- 1) группа 5 – фрагменты 400...1800 знаков с шагом 200 (655 фрагментов);
- 2) группа 6 – 2000...10 000 знаков с шагом 2000 (492 фрагмента);
- 3) группа 7 – 30 000...110 000 знаков с шагом 20 000 (300 фрагментов);
- 4) группа 8 – 100 000...350 000 знаков с шагом 50 000 (120 фрагментов).

Все фрагменты приведены к алфавиту из 31 буквы; пробелы и небуквенные символы удалены.

5.1. СРАВНЕНИЕ ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ БИГРАММ И ТРИГРАММ

Для оценки эффективности триграммного подхода проведено сравнение с методом Якобсена на биграммах на одних и тех же текстах выборки.

Использовалась ошибка полной идентификации O_1 – доля фрагментов, содержащих хотя бы один неправильно идентифицированный символ, относительно всех фрагментов данной длины [2].

На рис. 10 показана зависимость O_1 от объема текста для биграммного метода [2] и триграммного метода с разными вариантами нормировки.

На рис. 10 видно, что на малых объемах все методы дают высокие значения O_1 , но триграммный метод показывает несколько меньшую ошибку. Начиная с 800 символов триграммный метод с оптимальной нормировкой заметно превосходит биграммный: при 2000 символов O_1 снижается в 6,5 раза. С 6000 символов триграммный метод с оптимальной нормировкой достигает нулевой ошибки O_1 , нормировка к 15 000 обеспечивает наименьшую долю

ошибочных фрагментов на всём диапазоне. Нормировка к 5000 на малых и средних объемах близка к оптимальной, но свыше 6000 символов сохраняет 2...6 % ошибочных фрагментов.

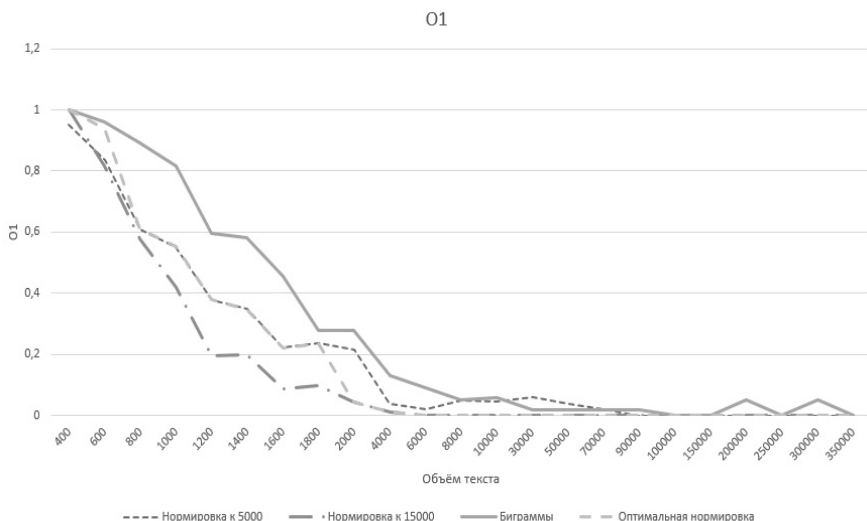


Рис. 8. График ошибок O1

Fig. 8. Error graph O1

Для оценки погрешности идентификации символов использовалась ошибка O2 [2] – среднее число неправильно идентифицированных знаков на один ошибочный фрагмент, нормированное на размер алфавита. Величина O2 вычислялась по формуле (9).

$$O2 = \frac{M}{N \cdot 31}, \quad (9)$$

где M – суммарное количество ошибочно идентифицированных символов для всех текстов данного объема; N – количество ошибочных текстов данного объема; 31 – размер алфавита.

На рис. 11 представлен график зависимости O2 от объема текста для биграммного [2] и триграммного методов с различными вариантами нормировки.

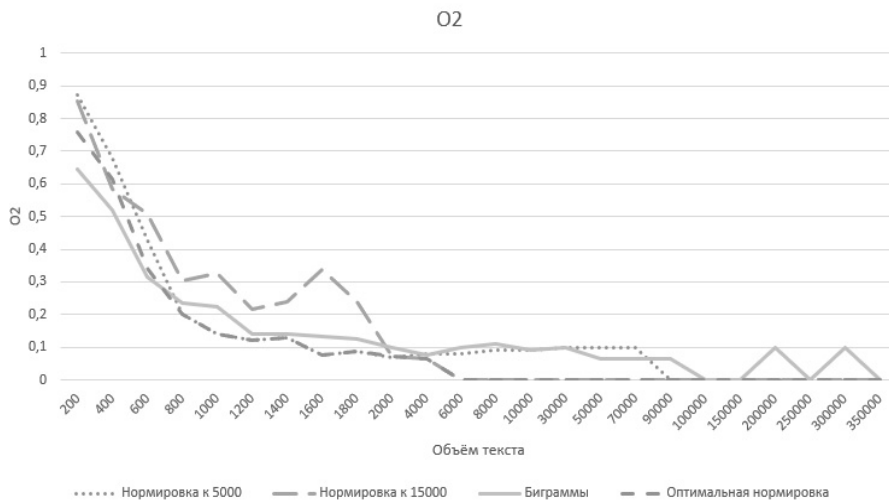


Рис. 9. График распределения ошибок O2

Fig. 9. Graph of O2 error distribution

Из рис. 11 видно, что в диапазоне 200...600 символов ошибка O2 для биграммного метода меньше на 8...16 %, чем для триграммного с оптимальной нормировкой. С 800 до 6000 символов соотношение меняется: триграммный метод дает ошибку O2 на 20...30 % ниже. Начиная с 6000 символов триграммный метод с оптимальной нормировкой достигает нулевого уровня O2, тогда как у биграммного остаются ненулевые значения. Ошибка снижается до нуля лишь к 90 000 символам.

Несмотря на более высокие затраты памяти и времени, снижение ошибки идентификации делает триграммный подход привлекательным для криптоанализа шифра простой замены при работе с текстами объемом более 800 символов.

6. МЕСТО РАБОТЫ В КОМПЛЕКСНОЙ ЗАЩИТЕ ОБЪЕКТА ИНФОРМАТИЗАЦИИ

Объект информатизации – совокупность информационных ресурсов, средств и систем обработки информации, а также помещений и объектов, в которых они установлены или которые предназначены для конфиденциальных переговоров [9]. Его защита обеспечивается комплексом мер: правовой,

технической, криптографической и физической [10]. Каждый вид защиты решает свой круг задач и по отдельности недостаточен.

Криптографическая подсистема обеспечивает конфиденциальность и целостность информации с помощью алгоритмов шифрования в СКЗИ. Стойкость алгоритмов подтверждается методами криптоанализа [11], включая статистические – анализ распределения символов шифротекста и тесты на случайность. Чем ближе распределение к равновероятному, тем выше стойкость [12]. Шифр простой замены по современным меркам недостаточно стойкий [11]: энтропия языка сохраняется в шифротексте, и статистические закономерности остаются доступны. Поэтому он служит учебной моделью статистического криптоанализа. Сформированный эталон частот триграмм и адаптированный метод Якобсена позволяют исследовать уязвимость шифра простой замены, оценить минимальный объем шифротекста для результативной атаки и сравнить триграммный и биграммный подходы. Триграммы применяются и в других задачах информационной безопасности. В работе [13] на фрагменте художественной литературы ошибка идентификации автора методом триграмм составила 23,8 % против 34,2 % для биграмм, подтверждая их большую информативность. Эталон частот триграмм – универсальный инструмент статистического анализа русскоязычных текстов.

ЗАКЛЮЧЕНИЕ

В работе решена задача адаптации метода Якобсена для криптоанализа шифра простой замены с использованием знаковых триграмм применительно к русскоязычным текстам. Основные результаты:

- сформирован эталон частот триграмм на основе репрезентативной выборки текстов общим объемом более 26 млн символов. Установлено, что состав топ-20 наиболее частотных триграмм стабилен для различных групп текстов, а группа текстов 3 (10 000...110 000 символов) покрывает 94,5 % эталонного разнообразия, что подтверждает достижение насыщения статистических характеристик языка;
- адаптирован метод Якобсена для триграммных распределений, и разработана его программная реализация. Экспериментально определена оптимальная стратегия нормировки частот в зависимости от объема анализируемого текста, позволяющая минимизировать ошибку идентификации знаков;
- проведен сравнительный анализ с биграммной реализацией метода. Триграммный подход обеспечивает существенно меньшую ошибку идентификации: для текстов объемом свыше 6000 символов ошибка O1 снижается до нуля, тогда как биграммный метод в этом диапазоне сохраняет ненулевые

значения вплоть до 90 000 символов. Ошибка O2 для триграммного метода в диапазоне 800...6000 символов на 20...30 % ниже по сравнению с биграммным подходом;

- количественно оценен минимальный объем шифротекста (около 6000 символов), при котором атака на основе триграммного эталона становится результативной, что подтверждает практическую ценность разработанного подхода для статистического криптоанализа русскоязычных текстов.

СПИСОК ЛИТЕРАТУРЫ

1. *Sabonchi A.K.S., Akay B.* A survey on the metaheuristics for cryptanalysis of substitution and transposition ciphers // *Computer Systems Science & Engineering*. – 2021. – Vol. 39 (1). – P. 87–104. – DOI: 10.32604/csse.2021.05365.
2. *Котов Ю.А.* Аппроксимация распределений частот буквенных биграмм текста для идентификации букв // *Труды СПИИРАН*. – 2017. – № 1 (50). – С. 190–208. – EDN XVLOWN.
3. *Васильев Е.М., Гусев К.Ю.* Анализ избыточности русскоязычного текста // *Вестник Воронежского государственного технического университета*. – 2010. – Т. 6, № 8. – С. 101–104.
4. *Котов Ю.А., Рассказов А.Д., Таронов Р.П.* Алгоритм идентификации простой замены букв текста по словарю // *Научные вести*. – 2019. – № 6 (11). – С. 201–212. – EDN SVJULP.
5. *Холоденко А.Б.* О построении статистических языковых моделей для систем распознавания русской речи // *Интеллектуальные системы*. – 2001. – Т. 6, № 1–4. – С. 381–394. – EDN YWAXTV.
6. *Munzner T.* Visualization analysis and design. – Boca Raton: CRC Press, 2015. – 404 p. – ISBN 978-1-4665-0893-4.
7. Алгоритм применения N-грамм для корректировки результатов распознавания / Т.В. Манжиков, О.А. Славин, И.А. Фараджев, И.М. Янишевский // *Вестник Российского фонда фундаментальных исследований*. – 2016. – № 4 (92). – С. 116–123. – DOI: 10.22204/2410-4639-2016-092-04-116-123.
8. *Котов Ю.А.* Криптографические методы защиты информации: шифры: учебное пособие. – Новосибирск: Изд-во НГТУ, 2016. – 60 с.
9. ГОСТ Р 51275–2006. Защита информации. Объект информатизации. Факторы, воздействующие на информацию. Общие положения. – М.: Стандартинформ, 2007.
10. ГОСТ Р 50922–2006. Защита информации. Основные термины и определения. – М.: Стандартинформ, 2008.

11. *Карпов А.В., Ишимуратов Р.А.* Введение в криптографию: учебное пособие. – Казань: Казан. ун-т, 2024. – 128 с.

12. Оценка стойкости криптографических шифров с помощью анализа шифртекстов / Т.С. Байшоланов, Ж.М. Алимжанова, Н. Байшолан, К.Е. Кубаев, К.С. Байшоланова // Известия Национальной академии наук Республики Казахстан. Серия физико-математическая. – 2022. – № 1. – С. 26–33. – DOI: 10.32014/2022.2518-1726.113.

13. *Кислицына М.Ю., Орлов Ю.Н.* Статистический анализ полного корпуса художественной литературы на русском языке и распознавание автора // Препринты ИПМ им. М.В. Келдыша. – 2024. – № 17. – С. 1–24. – EDN NIPYEE. – DOI: 10.20948/prepr-2024-17.

Котов Юрий Алексеевич, кандидат физико-математических наук, доцент кафедры защиты информации. E-mail: kotov@corp.nstu.ru

Башикиров Александр Петрович, лаборант кафедры защиты информации Новосибирского государственного технического университета. E-mail: kaf_zi@corp.nstu.ru

DOI: 10.17212/2782-2230-2026-3-41-63

Implementation of cryptanalysis of a simple substitution cipher using sign-based trigrams*

Yu.A. Kotov¹, A.P. Bashkirov²

¹*Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Information Security. E-mail: kotov@corp.nstu.ru*

²*Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, laboratory assistant of the Department of Information Security. E-mail: kaf_zi@corp.nstu.ru*

This paper examines the adaptation of the Jacobsen method-using sign-based trigrams-for the cryptanalysis of simple substitution ciphers applied to Russian-language texts. The study evaluates the robustness of the reference model by assessing the stability of the frequencies and ranks of the most common trigrams, the coverage of the set of non-zero trigrams, and the geometric similarity of their distribution across the cipher segments. Normalization was applied to ensure frequency comparability, and the optimal normalization strategy was experi-

* Received 14 July 2026.

mentally determined based on the volume of the text being analyzed. The method was implemented in software and tested on an independent dataset comprising popular science and literary texts. A comparative analysis with a bigram-based implementation demonstrated that the trigram approach yields a significantly lower identification error rate-dropping to zero for texts exceeding 6,000 characters-while maintaining acceptable computational costs. These results allow for a quantitative assessment of the minimum ciphertext volume required for a successful trigram-based attack and confirm the practical value of the trigram reference model for the statistical cryptanalysis of Russian-language texts.

Keywords: trigrams, frequency analysis, simple substitution cipher, Jacobsen's method, frequency reference, normalization

REFERENCES

1. Sabonchi A.K.S., Akay B. A survey on the metaheuristics for cryptanalysis of substitution and transposition ciphers. *Computer Systems Science & Engineering*, 2021, vol. 39 (1), pp. 87–104. DOI: 10.32604/csse.2021.05365.
2. Kotov Yu.A. Approximatsiya raspredelenii chastot bukvennykh bigramm teksta dlya identifikatsii bukv [Approximation of distributions of text characters bigrams frequencies for alphabetic characters identification]. *Trudy SPIIRAN = SPIIRAS Proceedings*, 2017, no. 1 (50), pp. 190–208.
3. Vasilyev E.M., Gusev K.Yu. Analiz izbytochnosti russkoyazychnogo teksta [The analysis of redundancy the Russian-speaking text]. *Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta = Bulletin of Voronezh State Technical University*, 2010, vol. 6, no. 8, pp. 101–104.
4. Kotov Yu.A., Rasskazov A.D., Taronov R.P. Algoritm identifikatsii prostoi zameny bukv teksta po slovaryu [Algorithm of identification one-to-one substitution text characters on dictionary]. *Nauchnye vesti = Scientific News*, 2019, no. 6 (11), pp. 201–212.
5. Kholodenko A.B. O postroenii statisticheskikh yazykovykh modelei dlya sistem raspoznavaniya russkoi rechi [On the construction of statistical language models for Russian speech recognition systems]. *Intellektual'nye sistemy = Intelligent Systems*, 2001, vol. 6, no. 1–4, pp. 381–394.
6. Munzner T. *Visualization analysis and design*. Boca Raton, CRC Press, 2015. 404 p. ISBN 978-1-4665-0893-4.
7. Manzhikov T.V., Slavin O.A., Faradzhev I.A., Janiszewski I.M. Algoritm primeneniya N-gramm dlya korrekcirovki rezul'tatov raspoznavaniya [N-grams algorithm application for the correction of recognition results]. *Vestnik Rossiiskogo fonda fundamental'nykh issledovanii = Russian Foundation for Basic Research Journal*, 2016, no. 4 (92), pp. 116–123. DOI: 10.22204/2410-4639-2016-092-04-116-123.

8. Kotov Yu.A. *Kriptograficheskie metody zashchity informatsii. Shifry* [Cryptographic methods of information protection. Ciphers]. Novosibirsk, NSTU Publ., 2016. 60 p.
9. GOST R 51275–2006. *Zashchita informatsii. Ob"ekt informatizatsii. Faktory, vozdeistvuyushchie na informatsiyu. Obshchie polozheniya* [State standard R 51275–2006. Protection of information. Object of informatisation. Factors influencing the information. General principles]. Moscow, Standartinform Publ., 2007.
10. GOST R 50922–2006. *Zashchita informatsii. Osnovnye terminy i opredeleniya* [State standard R 50922–2006. Protection of information. Basic terms and definitions]. Moscow, Standartinform Publ., 2008.
11. Karpov A.V., Ishmuratov R.A. *Vvedenie v kriptografiyu* [Introduction to cryptography]. Kazan, Kazan University Publ., 2024. 128 p.
12. Baisholanov T., Alimzhanova Zh., Baisholan N., Kubaev K., Baisholanova K. Otsenka stoikosti kriptograficheskikh shifrov s pomoshch'yu analiza shifrttekstov [Evaluation of the strength of cryptographic ciphers using ciphertext analysis]. *Izvestiya Natsional'noi akademii nauk Respubliki Kazakhstan. Seriya fiziko-matematicheskaya = News of the Academy of Sciences of the Republic of Kazakhstan. Series Physico-Mathematical*, 2022, no. 1, pp. 26–33. DOI: 10.32014/2022.2518-1726.113.
13. Kislitsyna M.Yu., Orlov Yu.N. Statisticheskii analiz polnogo korpusa khudozhestvennoi literatury na russkom yazyke i raspoznavanie avtora [Statistical analysis of the complete corpus of fiction in Russian and recognition of the author]. *Preprinty IPM im. M.V. Keldysha = Keldysh Institute preprints*, 2024, no. 17, pp. 1–24. DOI: 10.20948/prepr-2024-17.

Для цитирования:

Котов Ю.А., Башкиров А.П. Реализация криптоанализа шифра простой замены с использованием знаковых триграмм // Безопасность цифровых технологий. – 2026. – № 3 (122). – С. 41–63. – DOI: 10.17212/2782-2230-2026-3-41-63.

For citation:

Kotov Yu.A., Bashkirov A.P. Realizatsiya kriptoolnalyza shifra prostoi zameny s ispol'zovaniem znakovykh trigramm [Implementation of cryptanalysis of a simple substitution cipher using sign-based trigrams]. *Bezopasnost' tsifrovyykh tekhnologii = Digital Technology Security*, 2026, no. 3 (122), pp. 41–63. DOI: 10.17212/2782-2230-2026-3-41-63.

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

УДК 004

DOI: 10.17212/2782-2230-2026-3-64-83

**РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
ДЛЯ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ
В ИСПДн ВITRIX24***

Т.А. СИНЮШКИН¹, Г.А. ИСАЕВ²

¹ РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, лаборант кафедры защиты информации. E-mail: kaf_zi@corp.nstu.ru

² РФ, 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, старший преподаватель кафедры защиты информации. E-mail: g.isaev@corp.nstu.ru

В статье рассматривается разработка программного обеспечения для обезличивания персональных данных в информационной системе персональных данных (ИСПДн) Bitrix24. Актуальность работы обусловлена необходимостью использования данных CRM-систем для тестирования, аналитики и обучения при одновременном соблюдении требований Федерального закона № 152-ФЗ и Постановления Правительства РФ № 1154. Предложен адаптивный подход, основанный на классификации атрибутов CRM-сущностей по четырем категориям: прямые идентификаторы, квазиидентификаторы, чувствительные и служебные атрибуты. Разработан алгоритм динамического выбора методов обезличивания (маскирование, псевдонимизация, обобщение, подавление) в зависимости от категории поля и оцененного риска повторной идентификации. Представлена модульная архитектура программного обеспечения, включающая интеграционный модуль с Bitrix24, модуль обезличивания и ETL-подсистему. Экспериментальная оценка на синтетических данных показала снижение риска повторной идентификации с 70 до 22,5 % при сохранении 100 % логических связей между сущностями.

Ключевые слова: персональные данные, обезличивание, Bitrix24, CRM, ИСПДн, квазиидентификаторы, повторная идентификация, псевдонимизация, маскирование, ETL, API, защита информации

* Статья получена 16 июля 2026 г.

ВВЕДЕНИЕ

Актуальность темы обусловлена тем, что в корпоративной практике часто возникает необходимость использовать данные CRM-систем для тестирования программного обеспечения, построения отчетности, аналитики и обучения сотрудников. При этом применение реальных персональных данных в таких сценариях создает риск их неправомерного раскрытия, утечки или повторной идентификации субъектов. Полное удаление всех значимых сведений также не является оптимальным решением, поскольку в этом случае данные теряют прикладную ценность и не могут использоваться для анализа бизнес-процессов или проверки работы программных систем.

Особую сложность представляет обезличивание данных в CRM-системах, поскольку такие данные имеют связанную структуру. В Bitrix24 информация о контактах, компаниях, сделках, лидах, задачах и комментариях логически взаимосвязана. Персональные данные могут находиться не только в стандартных полях, но и в пользовательских атрибутах, текстовых комментариях и описаниях. Даже после удаления прямых идентификаторов сохраняется риск повторной идентификации субъекта по комбинации квазиидентификаторов.

Проблема исследования заключается в необходимости разработки подхода и программного решения для адаптивного обезличивания персональных данных в системе Bitrix24, обеспечивающего снижение риска повторной идентификации при сохранении структуры и прикладной ценности CRM-данных.

Целью настоящей работы является разработка программного обеспечения для обезличивания персональных данных в системе Bitrix24, обеспечивающего снижение риска повторной идентификации субъектов при сохранении полезности и логической структуры CRM-данных.

1. НОРМАТИВНАЯ БАЗА ОБРАБОТКИ И ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ В ИСПДн

1.1. ПОНЯТИЕ ПЕРСОНАЛЬНЫХ ДАННЫХ И ИХ КЛАССИФИКАЦИЯ

В соответствии с Федеральным законом № 152-ФЗ персональные данные определяются как любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу.

С научной точки зрения персональные данные классифицируются по степени их идентифицируемости. Выделяются идентификаторы, квазиидентифи-

каторы и нечувствительные атрибуты [1, 2]. Идентифицирующие данные позволяют однозначно установить личность субъекта (фамилия, имя, номер телефона). Квазиидентификаторы не обеспечивают однозначной идентификации при изолированном рассмотрении, однако их комбинация может привести к установлению личности (дата рождения, место проживания, род деятельности). Третью категорию составляют данные, не обладающие идентифицирующими свойствами даже при их комбинировании [2–4].

Понятие обезличивания персональных данных закреплено в Федеральном законе № 152-ФЗ и определяется как совокупность действий, в результате которых становится невозможным установить принадлежность данных конкретному субъекту без использования дополнительной информации.

1.2. ПОНЯТИЕ И СТРУКТУРА ИНФОРМАЦИОННЫХ СИСТЕМ ПЕРСОНАЛЬНЫХ ДАННЫХ (ИСПДн)

Информационная система персональных данных (ИСПДн) представляет собой совокупность технических, программных и организационных средств, обеспечивающих обработку персональных данных. Архитектура ИСПДн включает уровень данных, уровень приложений и уровень взаимодействия с пользователями.

Отдельным компонентом являются средства защиты информации, направленные на предотвращение несанкционированного доступа, утечки и модификации данных [5–7].

1.3. НОРМАТИВНО-ПРАВОВОЕ РЕГУЛИРОВАНИЕ ОБРАБОТКИ И ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Обработка персональных данных в информационных системах регулируется совокупностью нормативно-правовых актов, устанавливающих требования к их сбору, хранению, использованию и защите. В Российской Федерации ключевыми документами в данной области являются Федеральный закон № 152-ФЗ, Постановление Правительства РФ № 1119, Постановление Правительства РФ № 1154 и Приказ ФСТЭК России № 21, формирующие правовые и технические основы обработки персональных данных, включая требования к их защите и обезличиванию [8–11].

Постановление Правительства РФ № 1154 конкретизирует требования к процессу обезличивания, устанавливая необходимость обеспечения невозможности или существенного затруднения идентификации субъекта персональных данных. Существенным аспектом является требование учета риска повторной идентификации [10].

1.4. ТРЕБОВАНИЯ К ОБЕЗЛИЧИВАНИЮ ПЕРСОНАЛЬНЫХ ДАННЫХ В ИСПДн

Ключевым требованием является обеспечение баланса между уровнем защиты и сохранением полезности данных. Методы обезличивания должны выбираться таким образом, чтобы минимизировать потери информации при обеспечении требуемого уровня защиты.

Важным требованием является комплексный характер обработки данных. Необходимо учитывать все категории персональных данных, включая квази-идентификаторы, а также взаимосвязи между ними.

Дополнительным требованием является контекстно зависимый выбор методов обезличивания. Универсального подхода к обезличиванию персональных данных не существует, и поэтому выбор методов должен осуществляться с учетом конкретных условий обработки данных.

1.5. АНАЛИЗ BITRIX24 КАК ИНФОРМАЦИОННОЙ СИСТЕМЫ ПЕРСОНАЛЬНЫХ ДАННЫХ

Bitrix24 аккумулирует как идентифицирующие, так и поведенческие данные пользователей. Основными объектами обработки являются контакты, компании, сделки и лиды. Между указанными сущностями формируются логические связи.

Каждая из указанных сущностей может содержать поля, включающие персональные данные. Для сущности «контакт» характерны прямые идентификаторы (фамилия, имя, номер телефона, email), а также квазиидентификаторы (должность, регион, подразделение, дата рождения) [2, 3].

Особенностью CRM-систем является высокая степень связности данных, что существенно увеличивает риск повторной идентификации, поскольку позволяет восстанавливать информацию о субъекте на основе анализа связей и косвенных признаков.

Дополнительным фактором риска является возможность экспорта данных из системы, что повышает вероятность их несанкционированного использования.

2. АНАЛИЗ МЕТОДОВ ОБЕСПЕЧИВАНИЯ И ПОСТАНОВКА ЗАДАЧИ РАЗРАБОТКИ ПО

2.1. ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Одной из базовых моделей обезличивания является k -анонимность, предложенная Л. Суини. В рамках такого подхода набор данных считается

защищенным, если каждая запись неотличима как минимум от $k - 1$ других записей по набору квазиидентификаторов. Для достижения этого свойства используются операции обобщения и подавления значений. Несмотря на широкое распространение, k -анонимность обладает рядом ограничений, связанных с возможностью раскрытия чувствительной информации при недостаточном разнообразии данных [2].

С целью устранения указанных недостатков была предложена модель l -разнообразия, обеспечивающая наличие достаточного количества различных значений чувствительных атрибутов в каждой группе. Такой подход снижает риск атак, связанных с однородностью данных, однако не полностью устраняет возможность утечки информации [12].

Отдельное направление представляют методы дифференциальной приватности, основанные на добавлении случайного шума к результатам обработки данных. Такой подход обеспечивает формальные гарантии защиты, однако его применение ограничено в задачах, требующих сохранения точной структуры данных, например в системах класса CRM [13].

Наряду с формальными моделями, на практике применяются маскирование, псевдонимизация, агрегация и обобщение. Псевдонимизация занимает особое место, поскольку позволяет сохранить связность данных, но не исключает возможность восстановления исходной информации при наличии дополнительных данных.

Анализ показывает, что ни один из методов не обеспечивает универсального решения задачи обезличивания [14, 1].

2.2. АНАЛИЗ ПРИМЕНИМОСТИ МЕТОДОВ ОБЕЗЛИЧИВАНИЯ В ИСПДн BITRIX24

Bitrix24 характеризуется наличием разнородных данных, включающих идентифицирующие атрибуты, квазиидентификаторы и неструктурированные текстовые данные. Данные обладают высокой степенью связности, что существенно усложняет применение классических методов.

Маскирование может применяться для частичного сокрытия отдельных атрибутов, но не обеспечивает надежной защиты при атаках на основе сопоставления данных. Псевдонимизация сохраняет связность данных, но допускает восстановление исходных данных при наличии дополнительной информации. Методы обобщения эффективны для табличных данных, но ограничено применимы в системах со сложными взаимосвязями.

Применение отдельных методов в изолированном виде не позволяет обеспечить требуемый уровень защиты.

2.3. АНАЛИЗ РИСКОВ ПОВТОРНОЙ ИДЕНТИФИКАЦИИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Основным источником риска являются квазиидентификаторы. Классические исследования показывают, что 87 % населения США можно идентифицировать по комбинации почтового индекса, даты рождения и пола [3].

Риск повторной идентификации особенно высок в связанных данных. В CRM-системе один субъект может быть представлен в нескольких сущностях: контакт, сделка, задача, комментарий. Даже при удалении явных идентификаторов сохранение связей между сущностями и датами событий создает возможность для повторной идентификации на основе анализа временных паттернов и структуры взаимодействий.

2.4. ОГРАНИЧЕНИЯ И НЕДОСТАТКИ СУЩЕСТВУЮЩИХ РЕШЕНИЙ ОБЕЗЛИЧИВАНИЯ

Ключевым недостатком является ограниченная применимость существующих подходов. Большинство методов разрабатывались для обработки структурированных табличных данных и не учитывают сложные взаимосвязи между объектами.

Фундаментальной проблемой является компромисс между уровнем защиты и сохранением полезности данных. В существующих решениях этот компромисс часто не учитывается должным образом.

Серьезные ограничения связаны с обработкой неструктурированных данных. Большинство методов ориентированы на структурированные данные и не обеспечивают корректного выявления персональных сведений в текстах.

Отсутствие адаптивности методов также является существенным недостатком. В большинстве случаев применяются статические правила обработки, не учитывающие особенности конкретного набора данных и уровень риска.

2.5. ПОСТАНОВКА ЗАДАЧИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБЕЗЛИЧИВАНИЯ ДАННЫХ В BITRIX24

Целью настоящей работы является разработка программного обеспечения для обезличивания персональных данных в системе Bitrix24.

Формально задача может быть представлена следующим образом. Пусть задан исходный набор данных D , экспортированный из Bitrix24. Требуется построить функцию обезличивания F , в результате применения которой формируется обезличенный набор данных:

$$D' = F(D).$$

При этом преобразование должно удовлетворять следующим условиям:

$$R(D') = R_{\text{доп}},$$

$$U(D') = U_{\text{мин}},$$

$$S(D') = S(D),$$

где $R(D')$ – риск повторной идентификации субъекта персональных данных, $R_{\text{доп}}$ – допустимое пороговое значение риска, $U(D')$ – полезность обезличенных данных, $U_{\text{мин}}$ – минимально допустимый уровень полезности для решения прикладных задач, $S(D')$ – структура обезличенного набора данных и логические связи между сущностями Bitrix24.

Функция преобразования FFF должна учитывать типы данных и их классификацию, взаимосвязи между сущностями, контекст использования данных, а также требования нормативных документов. Необходимость учета квази-идентификаторов и риска повторной идентификации обусловлена тем, что даже после удаления прямых идентификаторов возможна идентификация субъекта на основе комбинации косвенных признаков и внешних источников данных [2–4].

Входными данными для разрабатываемой системы являются наборы персональных данных, полученные из Bitrix24 в структурированном и неструктурированном виде. Выходными данными являются обезличенные наборы данных, в которых риск идентификации субъектов снижен до допустимого уровня при сохранении аналитической и функциональной ценности данных.

К основным ограничениям задачи относятся:

- соответствие требованиям нормативных документов в области защиты персональных данных;
- необходимость обработки разнородных данных;
- сохранение целостности и связности данных;
- обеспечение масштабируемости обработки;
- соблюдение требований безопасности, установленных для ИСПДн [9, 11].

Ключевым элементом решения является необходимость оценки риска повторной идентификации. В соответствии с современными подходами этот риск рассматривается как вероятностная характеристика и должен учитываться при выборе методов обработки данных [14, 1].

3. РАЗРАБОТКА И РЕАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБЕЗЛИЧИВАНИЯ ДАННЫХ

3.1. ТРЕБОВАНИЯ К РАЗРАБАТЫВАЕМОМУ ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ

Программное обеспечение должно обеспечивать снижение риска повторной идентификации субъектов при сохранении полезности данных. Система должна учитывать особенности структуры данных Bitrix24, включая наличие идентифицирующих атрибутов, квазиидентификаторов и неструктурированных текстовых данных.

К основным функциональным требованиям системы относится поддержка процессов получения, обработки и сохранения данных, поступающих из Bitrix24. Программное обеспечение должно обеспечивать подключение к CRM-системе с использованием API Bitrix24, получение CRM-сущностей, извлечение метаданных полей и обработку экспортированных наборов данных. Поддержка интеграции с внешней системой является необходимым условием автоматизации процесса обезличивания и минимизации ручного вмешательства в обработку данных.

Система должна обеспечивать обработку основных сущностей Bitrix24, содержащих персональные данные (контакты, лиды, сделки и компании), и применять к ним набор правил обезличивания для стандартных и пользовательских полей.

Одним из ключевых функциональных требований является поддержка нескольких методов обезличивания персональных данных, таких как маскирование, удаление, псевдоминимизация, хэширование, генерация синтетических значений, обобщение данных.

Система должна обеспечивать автоматизированный запуск ETL-процессов, включающих этапы извлечения данных, их нормализации, преобразования и сохранения результатов. Для выполнения длительных операций обработки необходимо использование отдельного механизма оркестрации задач, позволяющего выполнять обработку асинхронно без блокировки пользовательских запросов.

Одним из обязательных требований является ведение журнала аудита действий пользователей и процессов обработки данных.

К нефункциональным требованиям относится обеспечение безопасности, расширяемости системы и воспроизводимости результатов обработки.

3.2. РАЗРАБОТКА МОДЕЛИ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Исходный набор данных представляется в виде множества записей:

$$D = \{d_1, d_2, \dots, d_n\},$$

где D – исходный набор данных, d_i – отдельная запись, n – количество записей. Каждая запись содержит набор атрибутов:

$$d_i = \{a_1, a_2, \dots, a_m\},$$

где a_j – отдельный атрибут записи, m – количество атрибутов.

В рамках модели множество атрибутов разделяется на четыре категории:

$$A = \{ID, QID, SA, SID\},$$

где ID – прямые идентификаторы, QID – квазиидентификаторы, SA – чувствительные атрибуты, SID – служебные и связующие идентификаторы.

Таблица 1

Table 1

Соответствие полей Bitrix24 категориям данных и методам обезличивания

Mapping of Bitrix24 fields to data categories and anonymization methods

Сущность Bitrix24	Поле	Категория	Метод обезличивания	Результат обработки
Контакт	ФИО	ID	Псевдоминимизация	Замена на условное значение, например «Клиент_001»
Контакт	Телефон	ID	Маскирование	Частичное сокрытие номера
Контакт	Email	ID	Защищенное хеширование или псевдонимизация	Исключение прямого восстановления адреса
Контакт	Дата рождения	QID	Обобщение	Замена точной даты возрастным диапазоном
Контакт	Город / регион	QID	Группировка	Замена точного значения укрупненной категорией

Окончание табл. 1

End of the Tab. 1

Сущность Bitrix24	Поле	Категория	Метод обезличивания	Результат обработки
Пользователь	Должность	<i>QID</i>	Обобщение	Замена конкретной должности группой должностей
Пользователь	Подразделение	<i>QID</i>	Группировка	Замена на более общий организационный уровень
Компания	Название компании	<i>QID / SID</i>	Псевдонимизация или обобщение	Замена на условное название или категорию
Сделка	Сумма сделки	<i>SA</i>	Обобщение	Замена точного значения диапазоном
Сделка	Коммерческие условия	<i>SA</i>	Подавление или маскирование	Соккрытие конфиденциальных фрагментов
Комментарий	Текст комментария	<i>SA</i>	Маскирование или подавление	Удаление или замена персональных фрагментов текста
CRM-сущность	ID контакта, сделки, компании	<i>SID</i>	Согласованная псевдонимизация	Сохранение связей между сущностями

Для прямых идентификаторов используется наиболее строгая обработка. Для квазиидентификаторов применяется снижение детализации (замена даты рождения возрастным диапазоном, города – регионом). Чувствительные атрибуты обрабатываются с учетом их значения для последующего анализа. Служебные идентификаторы не удаляются полностью, для них используется согласованная псевдонимизация.

3.3. РАЗРАБОТКА АЛГОРИТМА АДАПТИВНОГО ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Основная задача алгоритма адаптивного обезличивания персональных данных заключается в автоматизированной классификации полей CRM-сущностей Bitrix24, выборе подходящего метода обработки и формировании

обезличенного набора данных с сохранением логических связей между сущностями.

Классификация полей выполняется на основе анализа имени поля, его типа, метаданных и набора эвристических признаков. Для каждого поля вычисляется коэффициент уверенности классификации:

$$\text{conf}(f_i),$$

где f_i – анализируемое поле, $\text{conf}(f_i)$ – коэффициент уверенности в принадлежности поля к определенной категории данных. Чем больше количество совпадений с признаками определенной категории, тем выше значение данного коэффициента. Если коэффициент уверенности превышает заданное пороговое значение, полю назначается соответствующая категория.

Выбор метода осуществляется на основе категории данных и необходимости сохранения аналитической значимости поля. Преобразование данных выполняется последовательно для каждой записи исходного набора:

$$d'_i = T(d_i, p_j),$$

где d_i – исходная CRM-запись, p_j – выбранное правило обработки, d'_i – преобразованная запись.

После выполнения преобразования алгоритм проводит проверку результата, оценивая остаточный риск повторной идентификации. Если риск превышает допустимый уровень, алгоритм усиливает степень обезличивания.

Алгоритм предусматривает режим предварительного просмотра, позволяющий проверить корректность классификации полей до обработки всего набора данных.

3.4. ПРОЕКТИРОВАНИЕ АРХИТЕКТУРЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Архитектура разрабатываемой системы построена на модульном многоуровневом принципе, обеспечивающем безопасную обработку данных, разделение ответственности и возможность расширения. Выбранный подход позволяет отделить пользовательский интерфейс от прикладной логики, логику – от доступа к данным, а интеграцию с Bitrix24 – от алгоритмов обезличивания.

Т а б л и ц а 2

T a b l e 2

Основные компоненты архитектуры программного обеспечения**Key components of software architecture**

Компонент	Назначение
API-слой	Обработка HTTP-запросов, маршрутизация, валидация входных данных
Слой прикладной логики	Управление профилями обезличивания, правилами обработки, запуском задач и экспортом результатов
Модуль обезличивания	Реализация модели и алгоритма адаптивного обезличивания
Модуль интеграции с Bitrix24	Получение CRM-сущностей, метаданных и пользовательских полей из Bitrix24
ETL-подсистема	Выполнение длительных задач извлечения, преобразования и сохранения данных
Слой доступа к данным	Изоляция операций чтения и записи от прикладной логики
Хранилище данных	Хранение профилей, правил, параметров задач, журналов и обезличенных результатов
Подсистема аутентификации	Управление доступом пользователей к функциям системы
Подсистема аудита	Фиксация действий пользователей и этапов обработки без сохранения исходных персональных данных

С точки зрения безопасности архитектура предусматривает следующие меры:

- разграничение доступа пользователей по ролям;
- изоляцию модуля интеграции с Bitrix24 от остальных компонентов;
- отсутствие исходных персональных данных в системных логах;
- хранение параметров подключений в защищенном виде;
- журналирование операций обработки;
- выполнение длительных задач в отдельной ETL-подсистеме;
- сохранение только обезличенных результатов для дальнейшего анализа.

Выбранная архитектура обеспечивает модульность программного обеспечения, независимость компонентов, возможность масштабирования обработки данных и расширения набора методов обезличивания. Кроме того, она позволяет разделить обработку пользовательских запросов и длительные ETL-задачи, что повышает устойчивость системы при работе с большими объемами CRM-данных.

3.5. РЕАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Реализация программного обеспечения выполнена на языке Python с использованием асинхронного подхода и модульной структуры проекта. Серверная часть реализована на основе FastAPI [15]. Для описания структур данных применяются схемы Pydantic.

Модуль обезличивания реализует операции маскирования, псевдонимизации, обобщения, подавления и защищенного преобразования. Реализован модуль сопоставления полей для автоматического анализа стандартных и пользовательских атрибутов.

Интеграция с Bitrix24 реализована с использованием библиотеки httpx для асинхронных HTTP-запросов; ETL-задачи выполняются в Apache Airflow. Для разграничения доступа используется проверка JWT-токенов от Keycloak.

4. ОЦЕНКА ЭФФЕКТИВНОСТИ И РЕКОМЕНДАЦИИ

4.1. ПОДГОТОВКА ТЕКСТОВЫХ ДАННЫХ

В ходе разработки программного обеспечения был сформирован синтетический тестовый набор данных, моделирующий структуру CRM-данных Bitrix24. Набор включает основные CRM-сущности: контакты, лиды, сделки, компании – по 100 записей для каждой сущности. Данные содержат прямые идентификаторы, квазиидентификаторы, пользовательские поля и неструктурированные текстовые комментарии.

4.2. МЕТОДИКА ТЕСТИРОВАНИЯ И КРИТЕРИИ ОЦЕНКИ ЭФФЕКТИВНОСТИ

Оценка эффективности разрабатываемого программного обеспечения проводилась на основе сравнения двух состояний набора данных: исходного

и после обезличивания. Основным показателем являлось изменение риска повторной идентификации:

$$\Delta R = R(D) - R(D'),$$

где $R(D)$ – риск повторной идентификации в исходном наборе данных, $R(D')$ – риск повторной идентификации после обезличивания, ΔR – величина снижения риска.

Риск оценивается на основе анализа уникальности записей по квазиидентификаторам. Риск отдельной записи определяется следующим образом:

$$R_i = \frac{1}{|G_i|},$$

где R_i – риск повторной идентификации отдельной записи, G_i – группа записей с одинаковыми значениями квазиидентификаторов, $|G_i|$ – количество записей в данной группе.

Общий риск для набора данных рассчитывался как доля записей, отнесенных к группе повышенного риска:

$$R(G) = \frac{N_{risk}}{N_{total}} 100 \%,$$

где N_{risk} – количество записей с высоким риском повторной идентификации, N_{total} – общее количество записей в наборе данных.

После применения алгоритма аналогичный расчет выполнялся для обезличенного набора данных D' . Сравнение значений $R(D)$ и $R(D')$ позволяло определить степень снижения риска повторной идентификации.

4.3. РЕЗУЛЬТАТЫ РАБОТЫ АЛГОРИТМА ОБЕЗЛИЧИВАНИЯ

В результате применения разработанного алгоритма адаптивного обезличивания был сформирован обезличенный набор CRM-данных, включающий сведения по четырем основным сущностям Bitrix24: контактам, лидам, сделкам и компаниям. Обработка выполнялась на синтетическом тестовом наборе данных, подготовленном в разделе 4.1.

В ходе работы алгоритма было обработано 400 записей. Результаты обработки представлены в табл. 3. Стоит отметить, что применение согласованной псевдонимизации позволило сохранить все логические связи между CRM-сущностями.

Таблица 3

Table 3

Сводные результаты обработки текстовых данных**Summary results of text data processing**

CRM-сущность	Количество входных записей	Количество обработанных записей	Количество полей в сущности	Количество полей, подлежащих обезличиванию	Обработано полей	Ошибки
Контакты	100	100	36	19	19	0
Лиды	100	100	34	18	18	0
Сделки	100	100	32	14	14	0
Компании	100	100	30	12	12	0
Итого	400	400	132	63	63	0

4.4. ОЦЕНКА РИСКА ПОВТОРНОЙ ИДЕНТИФИКАЦИИ ПОСЛЕ ОБЕЗЛИЧИВАНИЯ

После завершения работы алгоритма обезличивания была проведена оценка риска повторной идентификации данных в соответствии с методикой, изложенной в разделе 4.2. Результаты оценки риска представлены в табл. 4.

Таблица 4

Table 4

Сравнение риска повторной идентификации до и после обезличивания**Comparison of re-identification risk before and after anonymization**

CRM-сущность	Количество записей	Риск до обработки $R(D)$, %	Риск после обработки $R(D')$, %	Снижение риска ΔR , %
Контакты	100	82	24	58
Лиды	100	78	26	52
Сделки	100	64	22	42
Компании	100	56	18	38
Итого	400	70	22,5	47,5

Наибольшее снижение риска наблюдается для сущностей, содержащих большое количество прямых идентификаторов и квазиидентификаторов.

Остаточный риск полностью не устраняется. Это связано с сохранением части структуры CRM-данных, возможной уникальностью отдельных комбинаций атрибутов и потенциальным использованием внешних источников информации.

4.5. ВЫВОДЫ ПО ЭФФЕКТИВНОСТИ РАЗРАБАТЫВАЕМОГО РЕШЕНИЯ И РЕКОМЕНДАЦИИ ПО ПРИМЕНЕНИЮ

Разработанное программное обеспечение выполняет автоматизированное обезличивание CRM-данных Bitrix24, снижает риск повторной идентификации и сохраняет структуру данных.

Практическая ценность заключается в возможности централизованного управления процессом обезличивания. Пользователь может настраивать профили обработки, изменять правила, выполнять предварительный просмотр и запускать ETL-процессы.

При применении разработанного ПО рекомендуется использовать отдельные профили для разных сценариев, выполнять предварительный просмотр, проверять пользовательские поля, ограничивать доступ к исходным данным и журналам, регулярно обновлять словари признаков ПДн.

Для дальнейшего развития целесообразно внедрение количественных метрик оценки риска, реализация проверки k -анонимности, расширение эвристик для пользовательских полей, развитие обработки неструктурированных текстовых данных, внедрение административного web-интерфейса.

ЗАКЛЮЧЕНИЕ

В статье представлен прототип программного обеспечения для адаптивного обезличивания персональных данных в системе Bitrix24, позволяющий снизить риск повторной идентификации при сохранении структуры и прикладной ценности CRM-данных. Разработанное решение поддерживает настройку профилей и правил обработки, классификацию стандартных и пользовательских полей, применение методов маскирования, псевдонимизации, обобщения, подавления и защищённого преобразования.

Результаты экспериментальной оценки подтверждают эффективность предложенного подхода: риск повторной идентификации снижен с 70 до 22,5 %, при этом алгоритм корректно обрабатывает различные категории персональных данных.

К числу ограничений прототипа относится отсутствие полной реализации формальных моделей k -анонимности, l -разнообразия и t -близости, а также зависимость качества обработки неструктурированных текстовых данных от полноты правил.

Перспективы дальнейшего развития включают внедрение формализованной проверки k -анонимности, расширение метрик оценки риска, совершенствование обработки текстовых данных и разработку административного веб-интерфейса.

СПИСОК ЛИТЕРАТУРЫ

1. Privacy-preserving data publishing: A survey of recent developments / B.C.M. Fung, K. Wang, R. Chen, P.S. Yu // *ACM Computing Surveys*. – 2010. – Vol. 42 (4). – P. 1–53.
2. Sweeney L. k -Anonymity: A model for protecting privacy // *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. – 2002. – Vol. 10 (5). – P. 557–570.
3. Sweeney L. Simple demographics often identify people uniquely. – Pittsburgh, 2000. – URL: <https://dataprivacylab.org/projects/identifiability/> (accessed: 08.09.2026).
4. Narayanan A., Shmatikov V. Robust De-anonymization of Large Sparse Datasets // *IEEE Symposium on Security and Privacy*. – IEEE, 2008. – P. 111–125.
5. Назаров Д.М., Саматов К.М. Основы обеспечения безопасности персональных данных. – Екатеринбург: УрФУ, 2019. – 216 с.
6. Обеспечение защиты персональных данных: методическое пособие / И.А. Баймакова, А.В. Новиков, А.И. Рогачев, А.Х. Хыдыров. – 3-е изд. – М.: 1С-Паблишинг, 2011. – 268 с.
7. Шаньгин В.Ф. Информационная безопасность компьютерных систем и сетей: учебное пособие. – М.: Форум, 2021. – 416 с.
8. О персональных данных: Федеральный закон РФ от 27.07.2006 № 152-ФЗ // *Российская газета*. – 2006. – 29 июля.
9. Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных: постановление Правительства РФ от 01.11.2012 № 1119 // *Собрание законодательства Российской Федерации*. – 2012. – № 45. – Ст. 6257.
10. Об утверждении требований к обезличиванию персональных данных, методов обезличивания персональных данных и Правил обезличивания персональных данных: постановление Правительства РФ от 01.08.2025 № 1154 // *Российская газета*. – 2025. – № 170. – С. 3.
11. Об утверждении состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных: приказ ФСТЭК России от 18.02.2013 № 21 // *Российская газета*. – 2013. – 17 мая.

12. L-diversity: Privacy beyond k-anonymity / A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam // ACM Transactions on Knowledge Discovery from Data. – 2007. – Vol. 1 (1). – P. 1–52.
13. Dwork C. Differential privacy // Automata, Languages and Programming. ICALP 2006. – Berlin: Springer, 2006. – P. 1–12. – (Lecture Notes in Computer Science; vol. 4052).
14. El Emam K. Guide to the de-Identification of personal health information. – Boca Raton: CRC Press, 2013. – 384 p.
15. FastAPI Documentation. – URL: <https://fastapi.tiangolo.com/> (accessed: 08.09.2026).

Синюшкин Тимофей Андреевич, лаборант кафедры защиты информации Новосибирского государственного технического университета. E-mail: kaf_zi@corp.nstu.ru

Исаев Глеб Андреевич, старший преподаватель кафедры защиты информации Новосибирского государственного технического университета. Основное направление научных исследований – булевы функции в теории кодирования и криптографии. E-mail: g.isaev@corp.nstu.ru

DOI: 10.17212/2782-2230-2026-3-64-83

Development of software for personal data anonymization in the Bitrix24 personal data information system*

T.A. Sinyushkin¹, G.A. Isaev²

¹ *Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, laboratory assistant at the Department of Information Security. E-mail: kaf_zi@corp.nstu.ru*

² *Novosibirsk State Technical University, 20 Karl Marx Prospekt, Novosibirsk, 630073, Russian Federation, Lecturer of the Department of Information Security. E-mail: g.isaev@corp.nstu.ru*

This article discusses the development of software for anonymizing personal data within the Bitrix24 personal data information system. The study is motivated by the need to utilize CRM system data for testing, analytics, and training purposes while simultaneously complying with the requirements of Federal Law No. 152-FZ and Russian Government Decree No. 1154. An adaptive approach is proposed, based on classifying CRM entity attributes into

* Received 16 July 2026.

four categories: direct identifiers, quasi-identifiers, sensitive attributes, and service attributes. An algorithm has been developed for the dynamic selection of anonymization methods (masking, pseudonymization, generalization, and suppression) based on field category and estimated re-identification risk. A modular software architecture is presented, comprising a Bitrix24 integration module, an anonymization module, and an ETL subsystem. Experimental evaluation using synthetic data demonstrated a reduction in re-identification risk from 70 % to 22.5 % while preserving 100 % of the logical relationships between entities.

Keywords: personal data, anonymization, Bitrix24, CRM, personal data information system, quasi-identifiers, re-identification, pseudonymization, masking, ETL, API, information protection

REFERENCES

1. Fung B.C.M., Wang K., Chen R., Yu P.S. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys*, 2010, vol. 42 (4), pp. 1–53.
2. Sweeney L. k-Anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002, vol. 10 (5), pp. 557–570.
3. Sweeney L. *Simple demographics often identify people uniquely*. Pittsburgh, 2000. Available at: <https://dataprivacylab.org/projects/identifiability/> (accessed 08.09.2026).
4. Narayanan A., Shmatikov V. Robust De-anonymization of Large Sparse Datasets. *IEEE Symposium on Security and Privacy*, 2008, pp. 111–125.
5. Nazarov D.M., Samatov K.M. *Osnovy obespecheniya bezopasnosti personal'nykh dannykh* [Fundamentals of personal data security]. Yekaterinburg, UrFU Publ., 2019. 216 p.
6. Baimakova I.A., Novikov A.V., Rogachev A.I., Khydyrov A.Kh. *Obespechenie zashchity personal'nykh dannykh* [Ensuring personal data protection]. 3rd ed. Moscow, 1C-Publishing, 2011. 268 p.
7. Shan'gin V.F. *Informatsionnaya bezopasnost' komp'yuternykh sistem i setei* [Information security of computer systems and networks]. Moscow, Forum Publ., 2021. 416 p.
8. On personal data: Federal Law of the Russian Federation No. 152-FZ dated July 27, 2006. *Rossiyskaya gazeta = Russia's Newspaper*, 2006, July 29. (In Russian).
9. On approving requirements for personal data protection during processing in personal data information systems: Decree of the Government of the Russian Federation No. 1119 dated November 1, 2012. *Sobranie zakonodatel'stva Rossiiskoi Federatsii = Collection of the legislation of the Russian Federation*, 2012, no. 45, Art. 6257. (In Russian).

10. On approval of requirements for the de-identification of personal data, methods for the de-identification of personal data, and Rules for the de-identification of personal data: Decree of the Government of the Russian Federation No. 1154 dated August 1, 2025. *Rossiyskaya gazeta = Russia's Newspaper*, 2025, no. 170, p. 3. (In Russian).
11. On approval of the composition and content of organizational and technical measures to ensure the security of personal data during processing in personal data information systems: Order of the FSTEC of Russia No. 21 dated February 18, 2013. *Rossiyskaya gazeta = Russia's Newspaper*, 2013, May 17. (In Russian).
12. Machanavajjhala A., Kifer D., Gehrke J., Venkatasubramanian M. L-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data*, 2007, vol. 1 (1), pp. 1–52.
13. Dwork C. Differential privacy. *Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science*, vol. 4052. Berlin, Springer, 2006, pp. 1–12.
14. El Emam K. *Guide to the de-identification of personal health information*. Boca Raton, CRC Press, 2013. 384 p.
15. *FastAPI documentation*. Available at: <https://fastapi.tiangolo.com/> (accessed 08.09.2026).

Для цитирования:

Синюшкин Т.А., Исаев Г.А. Разработка программного обеспечения для обезличивания персональных данных в ИСПДн Bitrix24 // Безопасность цифровых технологий. – 2026. – № 3 (122). – С. 64–83. – DOI: 10.17212/2782-2230-2026-3-64-83.

For citation:

Sinyushkin T.A., Isaev G.A. Razrabotka programmnoy obespecheniya dlya obezlichivaniya personal'nykh dannykh v ISPDn Bitrix24 [Development of software for personal data anonymization in the Bitrix24 personal data information system]. *Bezopasnost' tsifrovyykh tekhnologii = Digital Technology Security*, 2026, no. 3 (122), pp. 64–83. DOI: 10.17212/2782-2230-2026-3-64-83.

*МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ*

УДК 004

DOI: 10.17212/2782-2230-2026-3-84-98

**РАЗРАБОТКА МОДЕЛИ ПРОЦЕССА УПРАВЛЕНИЯ
ИНФРАСТРУКТУРНЫМИ УЯЗВИМОСТЯМИ
С УЧЕТОМ ИХ ПРИОРИТЕТА***

С.А. ГЛАДЫШЕВ¹, С.А. ГОЛУБЬ²

¹ РФ, 630008, г. Новосибирск, ул. Кирова, 23, ЗАО «Группа компаний С7», специалист по информационной безопасности. E-mail: s.galdyshev@s7.ru

² РФ, 630008, г. Новосибирск, ул. Кирова, 23, ЗАО «Группа компаний С7», специалист по информационной безопасности. E-mail: s.golub@s7.ru

В статье рассматриваются современные подходы к построению процесса управления уязвимостями в ИТ-инфраструктуре организаций с акцентом на приоритизацию уязвимостей и активов. Описываются ключевые этапы процесса: инвентаризация, выявление и оценка уязвимостей, приоритизация и устранение, а также мониторинг эффективности мер. Особое внимание уделено критериям классификации и приоритизации активов с учетом их бизнес-значимости, уровня доступа, последствий компрометации и сетевой доступности. Среди методов оценки уязвимостей анализируются существующие стандарты (CVSS 3.1), особенности zero day, доступность публичных эксплойтов и использование данных о трендах эксплуатации. Отмечается недостаточность стандартных методик приоритизации, предлагаются возможные усовершенствования, в том числе автоматизация с помощью современных сканеров и технологий машинного обучения. Разработанная модель приоритизации интегрируется в жизненный цикл процесса управления уязвимостями, закрепленный методическим документом ФСТЭК России, за счет чего формируется итоговая группировка из пяти групп приоритизации, возникающая на пересечении приоритета актива и критичности уязвимости и определяющая срок устранения, не превышающий регуляторный норматив. Эффективность разработанной модели оценивается с использованием отечественных нормативных метрик – операционных показателей методического документа ФСТЭК, коэффициента защищенности информации (КЗИ) и показателя зрелости процессов защиты информации (ПЗИ), что позволяет напрямую увязать результаты внедрения модели с требованиями регулятора. Полученные результаты могут быть использованы как готовое решение или в качестве основы для построения детализированного процесса управления уязвимостями с перспективой дальнейшей автоматизации.

* Статья получена 20 августа 2026 г.

Ключевые слова: управление уязвимостями, ИТ-инфраструктура, приоритизация уязвимостей, приоритизация активов, CVSS, уязвимости нулевого дня, доступность эксплойтов, оценка рисков, автоматизация безопасности, машинное обучение

ВВЕДЕНИЕ

Процесс управления уязвимостями является важнейшим аспектом кибербезопасности, который согласно [1] состоит из следующих этапов: мониторинга уязвимостей и оценки их применимости, оценки уязвимостей (определение уровня критичности), определения методов и приоритетов устранения, устранения уязвимостей.

Внедрение этого процесса в корпоративную инфраструктуру напрямую зависит от особенностей конечной инфраструктуры, а также информационных систем и средств защиты информации. Ключевой информационной системой процесса управления уязвимостями является сканер уязвимостей, парадигма использования которого рекомендуется вендорами с учетом технических возможностей предоставляемого решения. Так, Qualys реализует парадигму VMDR, объединяющую обнаружение активов, оценку, приоритизацию по бизнес-контексту и устранение в рамках единой платформы с непрерывным агентным сканированием [2]. Tenable предлагает гибридную модель – единый сканирующий движок Nessus, расширенный агентным и безагентным сбором данных, включая оценку облачных конфигураций [3]. Модуль Rapid7 в продукте InsightVM делает акцент на рискориентированной приоритизации через собственную метрику (VPR) и агентную непрерывную оценку конечных узлов без необходимости периодического сетевого сканирования [4]. В российской практике аналогичную парадигму предлагает Positive Technologies в продукте MaxPatrol VM, реализующем сбор данных об активах, выявление и приоритизацию уязвимостей на основе технологии Security Asset Management [5].

Однако предлагаемые методы построения процесса управления уязвимостями имеют ряд проблем, связанных с автоматизацией процесса, его внедрением в корпоративную инфраструктуру и приоритизацией выявленных уязвимостей. Большинство существующих решений автоматизируют лишь отдельные этапы процесса, а не полный цикл от обнаружения до подтвержденного устранения [6], а передача рутинных операций триажа автономным программным агентам пока не имеет достаточной эмпирической валидации [7]. Встраивание сканирования в CI/CD-конвейеры разработки требует согласования нескольких разнородных инструментов, что снижает эффект автоматизации на практике [8]. Формальные метрики приоритизации не учитывают контекст конкретной организации [9] и остаются внутренне противоречивыми: оценки одной и той же уязвимости в разных системах скоринга существенно

расходятся [10], единой общепринятой таксономии метрик приоритизации не выработано [11], что подтверждается масштабом всей базы NVD [12]. Наконец, эмпирические исследования показывают, что фактический уровень автоматизации в организациях заметно ниже декларируемого вендорами [13], социотехнические барьеры внедрения изучены меньше технических [14], а искусственный интеллект в реальном процессе управления уязвимостями выступает не нейтральным инструментом, а «соисполнителем» организационной рутины, изменяющим саму координацию команд [15], что в очередной раз подчеркивает потребность в использовании единой модели процесса с установленными метриками.

Основой задачей всех вышеупомянутых исследований помимо упрощения внедрения процесса управления уязвимостями в корпоративную инфраструктуру является повышение эффективности этого процесса. Этот показатель в отечественных компаниях оценивается через несколько взаимодополняющих показателей. На операционном уровне это соблюдение сроков устранения, установленных методическим документом ФСТЭК [1]: критический уровень – не более 24 часов, высокий – 7 дней, средний – 4 недели, низкий – 4 месяца, при этом сам уровень критичности определяется по отдельной методике оценки уязвимостей [16]. С вступлением в силу приказа ФСТЭК № 117 (с 01.03.2026) появляется формальный количественный норматив – коэффициент защищенности информации (КЗИ): значение от 0 до 1 рассчитывается не реже одного раза в полгода как взвешенная сумма 16 частных показателей по четырем группам мер, при этом устранение уязвимостей выделено в отдельный частный показатель (к32) в составе наиболее весомой группы «Защита информационных систем» [17]. Дополнительно тот же приказ вводит показатель зрелости процессов защиты информации (ПЗИ) – рассчитывается один раз в два года и оценивает не факт защищенности, а устойчивость и повторяемость самих процессов, включая управление уязвимостями [18].

Исходя из этого целью проводимого исследования является разработка модели процесса управления инфраструктурными уязвимостями, в основу которой ляжет метод приоритизации уязвимостей с учетом инфраструктурных особенностей. Для оценки эффективности разработанной модели будут использоваться отечественные метрики, закрепленные в соответствующих нормативных документах.

1. ОПИСАНИЕ МЕТОДА ПРИОРИТИЗАЦИИ УЯЗВИМОСТЕЙ

Для разработки эффективного метода приоритизации уязвимостей будут сформулированы стандартные критерии, свойственные каждой инфраструктуре с возможностью их адаптации. Целевой метод предполагает две основные

метрики приоритизации уязвимости: критичность актива, на котором находится уязвимость, и критичность самой уязвимости.

Приоритизацию активов информационной инфраструктуры предлагается осуществлять на основе критериев из табл. 1.

Таблица 1

Table 1

Критерии приоритизации активов информационной инфраструктуры

Prioritization criteria for information infrastructure assets

Критерий	Описание
Влияние на бизнес-процессы	Определяет, насколько отказ или компрометация актива приведет к прерыванию или остановке ключевых процессов организации
Уровень обрабатываемой или хранимой информации	Учитывается наличие в активе критически важных, конфиденциальных или персональных данных
Степень значимости в процессах инфраструктур	Оценивается, насколько актив необходим для функционирования других систем (например, является ли он центральным звеном или вспомогательным компонентом)
Доступ к управлению или конфигурации	Активы с административным доступом (например, контроллеры доменов) имеют более высокую критичность
Последствия компрометации или сбоя	Оценивается масштаб ущерба для организации: потеря данных, нарушение работы, финансовые, репутационные и правовые риски
Доступность резервирования и возможность быстрой замены	Если актив легко восстановить или заменить без значимых последствий, его критичность считается низкой
Сетевая доступность	Определяет уровень доступности актива из внешних и внутренних сетей: активы, находящиеся в открытом доступе для большого числа пользователей или доступные из внешних сетей (например, через интернет), обладают более высокой степенью критичности из-за повышенного риска атаки

На основе указанных в табл. 1 критериев формируется 3 основных приоритета активов: высокий, средний, низкий. К активам *высокого приоритета* относятся целевые и ключевые системы.

- При отказе или компрометации вызовут остановку/прерывание ключевых бизнес-процессов.
- Обработывают/хранят критически важную, особо конфиденциальную или персональную информацию.
- Являются центральным или основополагающим элементом инфраструктуры (точка отказа для многих систем).
- Обладают административным доступом либо средствами управления всей инфраструктурой.
- Компрометация данных влечет за собой максимальные негативные последствия: существенный финансовый, репутационный, правовой ущерб или нарушение регуляторных требований.

- Не подлежат быстрой замене или восстановлению без ущерба бизнесу.
- Доступны из внешних сетей или для большого числа пользователей.

Пример: контроллер домена, сервер баз данных с коммерческой тайной, шлюзовый сервер, сервер управления сетями, почтовый сервер компании.

Средний приоритет присваивается системам информационной безопасности и активам общего пользования.

- При отказе или компрометации оказывают влияние на бизнес-процессы, которые не приводят к полной остановке ключевой деятельности.
- Обработывают сведения средней чувствительности (например, служебную информацию, рабочие документы, общие корпоративные материалы).
- Являются важным компонентом, но не критической точкой отказа (например, поддерживающий сервер или функциональный сервис).
- Имеют ограниченные права управления или доступ к части инфраструктуры.
- Компрометация данных приведет к локальным потерям, нарушениям или необходимости восстановления.
- Могут быть быстро заменены/восстановлены с минимальным влиянием на бизнес-процессы.

- Сетевая доступность которых ограничена внутренней инфраструктурой или специально доверенными сегментами сети.

Пример: вспомогательные серверы, сервисы внутренней безопасности, рабочие станции ИБ-администраторов, серверы приложений внутри корпоративной сети.

К *низкому приоритету* относятся прочие активы: обычные рабочие станции, принтеры и другие периферийные устройства, сбой или потеря которых не приведет к значительным последствиям для организации. Однако стоит отметить, что распределение конкретных активов будет зависеть от организации конкретной инфраструктуры: тот или иной актив, указанный в примерах, может быть отнесен к другому уровню критичности. Перечень критериев является минимально достаточным, но может быть дополнен при необходимости.

Далее рассмотрим подход к приоритизации уязвимостей. Первым критерием критичности будет значение критичности классического подхода по распределению на категории: *critical*, *high*, *medium* и *low* по оценке CVSS 3.1. Одновременно стоит уделить внимание уязвимостям типа *zero day*, которым должен быть присвоен самый высокий приоритет при проработке. Под проработкой подразумевается оценка применимости в инфраструктуре, разработка и применение компенсирующих мер для исключения возможности эксплуатации уязвимости.

Также параметром оценки критичности уязвимости является совокупность способа ее эксплуатации (сетевой или не сетевой) и доступности актива (из внешней или не из внешней сети). Далее стоит обратить внимание на наличие публичного эксплойта для данной уязвимости. При сравнении двух уязвимостей с равным значением CVSS 3.1, находящихся на одном активе, та уязвимость, у которой опубликован эксплойт, будет иметь более высокий приоритет. Также стоит отметить значение конкретной уязвимости в трендовых списках. При наличии уязвимости в тренде наиболее эксплуатируемых уязвимостей (например, узнать эту информацию можно на ресурсе *cisa.kew*) необходимо повысить критичность ее уязвимости.

Без внимания не должны оставаться бюллетени и рекомендации регуляторов, в которых собраны актуальные уязвимости, способные повлечь за собой угрозу объектам критической информационной инфраструктуры. Наличие уязвимостей в информации от регуляторов должно служить фактором повышения критичности уязвимости, найденной в инфраструктуре.

В итоге все вышеперечисленные критерии приоритизации уязвимостей можно определить схемой, представленной на рисунке.

2. ОПИСАНИЕ ПРОЦЕССА УПРАВЛЕНИЯ УЯЗВИМОСТЯМИ С УЧЕТОМ МОДЕЛИ ПРИОРИТИЗАЦИИ

За основу модели процесса управления уязвимостями взят жизненный цикл процесса, закрепленный методическим документом ФСТЭК [1] и состоящий из пяти этапов: 1) мониторинг уязвимостей и оценка их применимости,

2) оценка уязвимостей, 3) определение методов и приоритетов устранения, 4) устранение уязвимостей и 5) контроль устранения. Встроим в этот цикл описанную выше модель приоритизации по активам (табл. 1) и по уязвимостям (рисунок) и рассмотрим, как это меняет содержание каждого этапа и какая группировка уязвимостей формируется в результате.

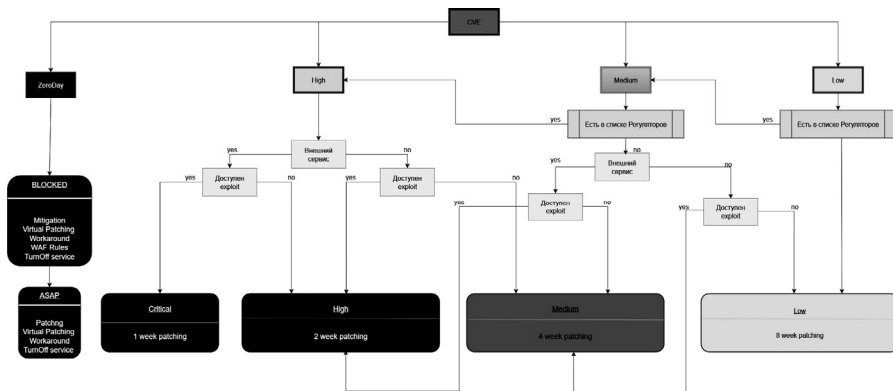


Схема приоритизации уязвимостей

Vulnerability prioritization scheme

На этапе мониторинга уязвимостей и оценки их применимости к стандартному сбору данных из БДУ ФСТЭК и иных источников уязвимостей добавляется процесс инвентаризации активов с присвоением каждому активу приоритета по критериям табл. 1 (высокий, средний, низкий). Это позволяет уже на входе связывать каждую обнаруженную уязвимость не только с ее собственными характеристиками, но и с критичностью актива, на котором она найдена.

На этапе оценки уязвимостей применяется схема, представленная на рисунке: уязвимость классифицируется по значению CVSS 3.1 (critical/high/medium/low), после чего эта базовая оценка корректируется тремя дополнительными факторами: 1) принадлежностью к уязвимостям типа zero day, 2) наличием уязвимости в бюллетенях регуляторов и 3) доступностью публичного эксплойта. Результатом этапа является итоговая категория критичности уязвимости, полученная не только из CVSS, но и с учетом реальной вероятности и актуальности эксплуатации.

Ключевое отличие описываемой модели реализовано на этапе определения методов и приоритетов устранения: именно здесь категория критичности

уязвимости (со второго этапа) пересекается с приоритетом актива (с первого этапа), формируя итоговую группу приоритизации с закреплённым сроком устранения. Правило пересечения формулируется следующим образом: базовым сроком устранения для каждого уровня критичности уязвимости служит норматив ФСТЭК [1] (критический – 24 часа, высокий – 7 дней, средний – 4 недели, низкий – 4 месяца); для уязвимостей на активах высокого приоритета этот срок ужесточается на одну ступень (высокий уровень критичности – 3 суток вместо 7 дней, средний – 7 дней вместо 4 недель, низкий – 4 недели вместо 4 месяцев), тогда как для активов среднего приоритета норматив ФСТЭК применяется без изменений, а для активов низкого приоритета уязвимости средней и низкой критичности устраняются в пределах того же норматива, но по остаточному принципу – после устранения уязвимостей той же категории критичности на более приоритетных активах. Ни при каких условиях итоговый срок не может превышать норматив ФСТЭК для соответствующего уровня критичности уязвимости: понижение приоритета актива не является основанием для нарушения регуляторных сроков, но может быть основанием для их ужесточения.

В результате пересечения трех уровней приоритета активов и четырёх категорий критичности уязвимостей формируется группировка из пяти итоговых групп приоритизации, представленная в табл. 2.

Таблица 2

Table 2

Итоговая группировка приоритетов устранения уязвимостей

Final grouping of vulnerability remediation priorities

Приоритет актива / Критичность уязвимости	Critical	High	Medium	Low
Высокий	Группа 1 (≤ 24 ч)	Группа 2 (≤ 3 сут.)	Группа 3 (≤ 7 дн.)	Группа 4 (≤ 4 нед.)
Средний	Группа 1 (≤ 24 ч)	Группа 3 (≤ 7 дн.)	Группа 4 (≤ 4 нед.)	Группа 5 (≤ 4 мес.)
Низкий	Группа 1 (≤ 24 ч)	Группа 3 (≤ 7 дн.)	Группа 5 (≤ 4 мес., в очереди последний)	Группа 5 (≤ 4 мес., в очереди последний)

Таким образом, вместо единственного основания для приоритизации (критичность уязвимости по CVSS), используемого в базовом варианте методики ФСТЭК, предложенная модель формирует пять групп приоритизации на пересечении критичности уязвимости и приоритета актива, на котором она обнаружена. Такая группировка не нарушает регуляторных нормативов сроков устранения [1] – итоговый срок для любой уязвимости не может быть больше официального норматива для ее уровня критичности, но позволяет обоснованно сокращать сроки для уязвимостей на наиболее значимых активах инфраструктуры и выстраивать очередность устранения внутри одной регуляторной группы сроков. Этапы устранения уязвимостей и контроля устранения (четвертый и пятый этапы цикла ФСТЭК) выполняются далее в соответствии с присвоенной группой. На этапе контроля собираются метрики эффективности процесса.

3. ОЦЕНКА ЭФФЕКТИВНОСТИ МОДЕЛИ ПРОЦЕССА УПРАВЛЕНИЯ УЯЗВИМОСТЯМИ

Эффективность предложенной модели оценивается метриками, закрепленными в отечественных нормативных документах.

На операционном уровне применимы показатели, прямо следующие из методического документа [1]: доля уязвимостей, устраненных в срок, установленный для присвоенной группы приоритизации (см. табл. 2); количество уязвимостей, не устраненных в срок, в разрезе групп; процент покрытия активов инфраструктуры сканером уязвимостей; количество обнаруженных и устраненных уязвимостей за выбранный период. Поскольку предложенная модель присваивает каждой уязвимости не только категорию критичности, но и конкретную группу с фиксированным сроком устранения, перечисленные показатели могут рассчитываться отдельно для каждой из пяти групп табл. 2, что дает более детальную картину эффективности процесса, чем усредненный показатель по всей организации.

На уровне формального норматива применим коэффициент защищенности информации (КЗИ), вводимый приказом ФСТЭК № 117 и методикой его расчета [17]. Устранение уязвимостей учитывается в КЗИ отдельным частным показателем «к32» в составе группы «Защита информационных систем» с наибольшим весовым коэффициентом ($R3 = 0,35$) среди всех групп показателей [17, 18]. Так как частный показатель «к32» обнуляется при повторном (в течение 12 месяцев) невыполнении соответствующей меры, а обнуление одного показателя обнуляет весь весовой коэффициент группы [17], для организации критически важно демонстрировать устойчивое соблюдение сроков устранения уязвимостей, а не разовое соответствие. Прямое соответствие структуры предложенной модели (группы приоритизации с фиксированными и не превышающими норматив ФСТЭК сроками) требованиям к частному

показателю «к32» позволяет использовать данные модели непосредственно как основание для расчета и подтверждения этого показателя при очередной (не реже одного раза в полгода) оценке КЗИ.

На уровне зрелости процессов применим показатель зрелости процессов защиты информации (ПЗИ), также вводимый приказом ФСТЭК № 117 [17, 18] и рассчитываемый не реже одного раза в два года. Показатель ПЗИ оценивает не факт защищенности, а устойчивость и повторяемость самих процессов защиты, включая управление уязвимостями. Формализация процесса приоритизации в виде фиксированных критериев (см. табл. 1), детерминированной схемы оценки уязвимостей (см. рисунок) и однозначной матрицы групп (см. табл. 2) переводит процесс управления уязвимостями из категории решений, принимаемых экспертом ситуативно, в категорию документированной и воспроизводимой процедуры, т. е. напрямую отвечает на тот критерий, по которому оценивается ПЗИ.

Представленные в этом разделе выводы носят методический характер: количественная оценка прироста значений КЗИ и ПЗИ, а также фактического сокращения доли просроченных уязвимостей при внедрении модели требует апробации на данных конкретной организации и является предметом дальнейшего исследования. При этом уже на уровне структуры модели видно, что все ее элементы: приоритизация активов, многофакторная оценка уязвимостей, группировка со сроками устранения – напрямую соответствуют перечню показателей, по которым эффективность процесса управления уязвимостями оценивается регулятором, что отличает предложенную модель от универсальных существующих на рынке решений.

ЗАКЛЮЧЕНИЕ

В настоящей статье описана модель приоритизации уязвимостей, показана ее интеграция в жизненный цикл процесса управления уязвимостями, закрепленный методическим документом ФСТЭК. Ключевым результатом этой интеграции стало формирование итоговой группировки из пяти групп приоритизации (см. табл. 2), возникающей на пересечении приоритета актива и критичности уязвимости. Такая группировка не нарушает регуляторных сроков устранения, установленных методическим документом ФСТЭК, но позволяет обоснованно ужесточать их для наиболее значимых активов инфраструктуры и выстраивать очередность устранения внутри одной регуляторной группы сроков, т. е. переводит приоритизацию с единственного основания (критичности уязвимости) на два взаимодополняющих основания одновременно.

Разработанная модель адаптивна: она может использоваться как готовое решение для широкого круга организаций или служить отправной точкой для построения специализированных процессов в сложных инфраструктурах с учетом уникальных особенностей. Модель позволяет добавлять в нее крите-

рии оценки критичности уязвимости, используя заданный вектор рассмотрения метрик CVSS 3.1 и иных факторов целевой инфраструктуры, влияющих на критичность обнаруженной уязвимости.

Внедрение предложенной модели позволяет структурировать процесс устранения уязвимостей, эффективно расставлять приоритеты и оптимально использовать ресурсы компании для снижения рисков. Таким образом, статья формирует основу для построения результативной системы управления инфраструктурными уязвимостями, отвечающей современным требованиям безопасности и нормативным требованиям ФСТЭК России.

СПИСОК ЛИТЕРАТУРЫ

1. ФСТЭК России. Руководство по управлению уязвимостями программного обеспечения и программно-аппаратных средств в органе (организации). Методический документ. – М., 2023. – 29 с. – URL: <https://base.garant.ru/406955550/> (дата обращения: 07.09.2026).
2. Qualys VMDR – Vulnerability management, detection and response (official documentation). – URL: https://docs.qualys.com/en/vm/latest/about_qualys_vm_vmdr.htm (accessed: 07.09.2026).
3. Tenable Nessus vs Qualys VMDR vs Rapid7 InsightVM: Vulnerability scanner comparison for enterprise security teams (2026). – URL: <https://www.decryptiondigest.com/blog/tenable-nessus-vs-qualys-vmdr-vs-rapid7-insightvm-vulnerability-scanner-comparison> (accessed: 07.09.2026).
4. Rapid7 InsightVM vs Tenable Nessus 2026: Coverage, risk. – <https://www.decryptiondigest.com/blog/rapid7-vs-tenable-vulnerability-management> (accessed: 07.09.2026).
5. MaxPatrol VM – Positive Technologies. – <https://global.ptsecurity.com/en/products/mp-vm/> (accessed: 07.09.2026).
6. SEEMA: An automation framework for vulnerability assesement and penetration testing / Y. Tyagi, S. Shekhar, P. Abhishek, S. Bhardwaj // 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN). – Vellore, India, 2023. – P. 1–5. – DOI: 10.1109/ViTECoN58111.2023.10157032.
7. Automated CVE Triage: Ai-agent framework for scalable vulnerability triaging and security automation / N. Amirov, E. Huseynov, N. Aliyev, Ş. Bahtiyar // 2025 18th International Conference on Information Security and Cryptology (ISCTürkiye). – Ankara, Türkiye, 2025. – P. 1–6. – DOI: 10.1109/ISCTrkiye68593.2025.11224852.
8. Cyril H.P., Kumara S. DevSecOps-driven security integration in the software development lifecycle using CI/CD pipelines // 2026 IEEE 5th International

Conference on AI in Cybersecurity (ICAIC). – Houston, TX, USA, 2026. – P. 1–6. – DOI: 10.1109/ICAIC67076.2026.11395737.

9. Jung B., Li Y., Bechor T. CAVP: A context-aware vulnerability prioritization model // Computers and Security. – 2022. – Vol. 116. – P. 102639. – DOI: 10.1016/j.cose.2022.102639.

10. Conflicting scores, confusing signals: An empirical study of vulnerability scoring systems / V. Koscinski, M. Nelson, A. Okutan, R. Falso, M. Mirakhorli // Proceedings of 32nd ACM Conference on Computer and Communications Security (CCS) (ACM CCS 2025). – Taipei, Taiwan, 2025. – URL: <https://arxiv.org/pdf/2508.13644> (accessed: 07.09.2026).

11. A survey on vulnerability prioritization: Taxonomy, metrics, and research challenges / Y. Jiang, N. Oo, Q. Meng, H.W. Lim, B. Sikdar // arXiv. – 2025. – URL: <https://arxiv.org/pdf/2502.11070> (accessed: 07.09.2026).

12. Fragmentation of CVSS scores in the NVD: A quantitative analysis of inconsistency across vulnerability scoring standards / M. Walkowski, M. Nowak, A. Balsam, K. Nowak, J. Oko, S. Sujecki // Computers and Security. – 2026. – Vol. 169. – P. 104978. – DOI: 10.1016/j.cose.2026.104978.

13. An empirical study of automation in software security patch management / N. Dissanayake, A. Jayatilaka, M. Zahedi, M.A. Babar // arXiv. – 2022. – URL: <https://arxiv.org/pdf/2209.01518> (accessed: 08.09.2026).

14. Software security patch management: A systematic literature review of challenges, approaches, tools and practices / N. Dissanayake, A. Jayatilaka, M. Zahedi, M.A. Babar // Information and Software Technology. – 2022. – Vol. 144. – DOI: 10.1016/j.infsof.2021.106771.

15. Saadallah M., Shahim A., Khapova S. Rethinking vulnerability management: How AI and automation reshape organizational routines and supports adaptive cybersecurity systems // Systems (MDPI). – 2026. – Vol. 14 (5). – P. 573. – DOI: 10.3390/systems14050573.

16. Методика оценки уровня критичности уязвимостей программных, программно-аппаратных средств: утв. ФСТЭК России 28.10.2022. – URL: <https://bazanpa.ru/fstek-rossii-metodika-ot28102022-h5766872/> (дата обращения: 08.09.2026).

17. Методика оценки показателя состояния технической защиты информации в информационных системах и обеспечения безопасности значимых объектов КИИ РФ: утв. ФСТЭК России 11.11.2025). – URL: <https://www.garant.ru/products/ipo/prime/doc/412956179/> (дата обращения: 08.09.2026).

18. Шелыгин И. Приказ ФСТЭК № 117: понятный разбор для бизнеса: КЗИ и ПЗИ – практическое значение показателей. – 2026. – URL: <https://passwork.ru/blog/prikaz-fstek-117/> (дата обращения: 08.09.2026).

Гладышев Сергей Александрович, специалист по информационной безопасности ЗАО «Группа компаний С7». Основное направление научных исследований – разработка автоматизированных сценариев моделирования атак информационной безопасности. E-mail: s.galdyshev@s7.ru

Голубь Сергей Андреевич, специалист по информационной безопасности ЗАО «Группа компаний С7». Основное направление научных исследований – разработка автоматизированных сценариев моделирования атак информационной безопасности. E-mail: s.golub@s7.ru

DOI: 10.17212/2782-2230-2026-3-84-98

Development of a process model for managing infrastructure vulnerabilities with regard to their priority*

S.A. Gladyshev¹, S.A. Golub²

¹ CJSC "Group of Companies S7", 23 Kirova Street, Novosibirsk, 630008, Russian Federation. Information Security Specialist. E-mail: s.galdyshev@s7.ru

² CJSC "Group of Companies S7", 23 Kirova Street, Novosibirsk, 630008, Russian Federation, Information Security Specialist. E-mail: s.golub@s7.ru

This article examines modern approaches to building a vulnerability management process within the IT infrastructure of organizations, with an emphasis on the prioritization of vulnerabilities and assets. The key stages of the process are described: inventory, vulnerability identification and assessment, prioritization and remediation, and monitoring the effectiveness of the measures taken. Particular attention is given to the criteria for classifying and prioritizing assets based on their business significance, access level, consequences of compromise, and network accessibility. Among the vulnerability assessment methods, the article analyzes existing standards (CVSS 3.1), the specifics of zero-day vulnerabilities, the availability of public exploits, and the use of exploitation trend data. The insufficiency of standard prioritization methods is noted, and possible improvements are proposed, including automation through modern scanners and machine learning technologies. The developed prioritization model is integrated into the vulnerability management lifecycle established by the methodological document of FSTEC of Russia, forming a resulting grouping of five prioritization groups derived from the intersection of asset priority and vulnerability criticality, each with a remediation deadline that does not exceed the regulatory requirement. The effectiveness of the developed model is assessed using domestic regulatory metrics – the operational indicators of the FSTEC methodological document, the information security coefficient (KZI), and the security process maturity indicator (PZI) – allowing the results of the model's implementation to be directly linked to regulatory requirements. The results obtained can be used as a ready-made solution or as a basis for building a more detailed vulnerability management process with a view to further automation.

Keywords: vulnerability management, IT infrastructure, vulnerability prioritization, asset prioritization, CVSS, zero-day vulnerabilities, exploit availability, risk assessment, security automation, machine learning

* Received 20 August 2026.

REFERENCES

1. FSTEC of Russia. *Guidelines for the management of vulnerabilities in software and hardware within a government body (organization)*. Methodological document, 2023. 29 p. (In Russian). Available at: <https://base.garant.ru/406955550/> (accessed 07.09.2026).
2. *Qualys VMDR – Vulnerability Management, Detection and Response* (official documentation). Available at: https://docs.qualys.com/en/vm/latest/about_qualys_vm_vmdr.htm (accessed 07.09.2026).
3. *Tenable Nessus vs Qualys VMDR vs Rapid7 InsightVM: Vulnerability scanner comparison for enterprise security teams* (2026). Available at: <https://www.decryptiondigest.com/blog/tenable-nessus-vs-qualys-vmdr-vs-rapid7-insightvm-vulnerability-scanner-comparison> (accessed 07.09.2026).
4. *Rapid7 InsightVM vs Tenable Nessus 2026: Coverage, risk*. Available at: <https://www.decryptiondigest.com/blog/rapid7-vs-tenable-vulnerability-management> (accessed 07.09.2026).
5. *MaxPatrol VM – Positive Technologies* (product page). Available at: <https://global.ptsecurity.com/en/products/mp-vm/> (accessed 07.09.2026).
6. Tyagi Y., Shekhar S., Abhishek P., Bhardwaj S. SEEMA: An automation framework for vulnerability assessment and penetration testing. *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, Vellore, India, 2023, pp. 1–5. DOI: 10.1109/ViTECoN58111.2023.10157032.
7. Amirov N., Huseynov E., Aliyev N., Bahtiyar Ş. Automated CVE Triage: Ai-agent framework for scalable vulnerability triaging and security automation. *2025 18th International Conference on Information Security and Cryptology (ISCTürkiye)*, Ankara, Türkiye, 2025, pp. 1–6. DOI: 10.1109/ISCTrkiye68593.2025.11224852.
8. Cyril H.P., Kumara S. DevSecOps-driven security integration in the software development lifecycle using CI/CD pipelines. *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, Houston, TX, USA, 2026, pp. 1–6. DOI: 10.1109/ICAIC67076.2026.11395737.
9. Jung B., Li Y., Bechor T. CAVP: A context-aware vulnerability prioritization model. *Computers and Security*, 2022, vol. 116, p. 102639. DOI: 10.1016/j.cose.2022.102639.
10. Koscinski V., Nelson M., Okutan A., Falso R., Mirakhorli M. Conflicting scores, confusing signals: An empirical study of vulnerability scoring systems. *Proceedings of 32nd ACM Conference on Computer and Communications Security (CCS) (ACM CCS 2025)*, Taipei, Taiwan, 2025. Available at: <https://arxiv.org/pdf/2508.13644> (accessed 07.09.2026).
11. Jiang Y., Oo N., Meng Q., Lim H.W., Sikdar B. A survey on vulnerability prioritization: Taxonomy, metrics, and research challenges. *arXiv*, 2025. Available at: <https://arxiv.org/pdf/2502.11070> (accessed 07.09.2026).

12. Walkowski M., Nowak M., Balsam A., Nowak K., Oko J., Sujecki S. Fragmentation of CVSS scores in the NVD: A quantitative analysis of inconsistency across vulnerability scoring standards. *Computers and Security*, 2026, vol. 169, p. 104978. DOI: 10.1016/j.cose.2026.104978.

13. Dissanayake N., Jayatilaka A., Zahedi M., Babar M.A. An empirical study of automation in software security patch management. *arXiv*, 2022. Available at: <https://arxiv.org/pdf/2209.01518> (accessed 08.09.2026).

14. Dissanayake N., Jayatilaka A., Zahedi M., Babar M.A. Software security patch management: A systematic literature review of challenges, approaches, tools and practices. *Information and Software Technology*, 2022, vol. 144. DOI: 10.1016/j.infsof.2021.106771.

15. Saadallah M., Shahim A., Khapova S. Rethinking vulnerability management: How AI and automation reshape organizational routines and supports adaptive cybersecurity systems. *Systems (MDPI)*, 2026, vol. 14 (5), p. 573. DOI: 10.3390/systems14050573.

16. *Methodology for assessing the criticality level of vulnerabilities in software and hardware/firmware* (approved by FSTEC of Russia, 28.10.2022). (In Russian). Available at: <https://bazanpa.ru/fstek-rossii-metodika-ot28102022-h5766872/> (accessed 08.09.2026).

17. *Methodology for assessing the indicator of the state of technical information protection in information systems and ensuring the security of significant facilities of critical information infrastructure of the Russian Federation* (approved by FSTEC of Russia, 11.11.2025). (In Russian). Available at: <https://www.garant.ru/products/ipo/prime/doc/412956179/> (accessed 08.09.2026).

18. Shelygin I. *Prikaz FSTEC № 117: ponyatnyi razbor dlya biznesa: KZI i PZI – prakticheskoe znachenie pokazatelei* [Overview of FSTEC Order No. 117: KZI (Information Security Coefficient) and PZI (Security Process Maturity Indicator) – practical significance of the indicators]. 2026. Available at: <https://passwork.ru/blog/prikaz-fstek-117/> (accessed 08.09.2026).

Для цитирования:

Гладышев С.А., Голубь С.А. Разработка модели процесса управления инфраструктурными уязвимостями с учетом их приоритета // Безопасность цифровых технологий. – 2026. – № 3 (122). – С. 84–98. – DOI: 10.17212/2782-2230-2026-3-84-98.

For citation:

Gladyshev S.A., Golub C.A. Razrabotka modeli protsessu upravleniya infrastrukturnymi uyazvimostyami s uchetom ikh prioriteta [Development of a process model for managing infrastructure vulnerabilities with regard to their priority]. *Bezopasnost' tsifrovyykh tekhnologii = Digital Technology Security*, 2026, no. 3 (122), pp. 84–98. DOI: 10.17212/2782-2230-2026-3-84-98.

ПРАВИЛА ДЛЯ АВТОРОВ

УСЛОВИЯ ПРИЕМА СТАТЕЙ

Все статьи и сопровождающие их материалы в журнал подаются через сайт журнала в электронном виде после регистрации всех авторов статьи. Регистрация обязывает каждого автора иметь международный идентификационный номер ORCID. Иные варианты подачи материалов не рассматриваются.

Автор (один из соавторов) в своем личном кабинете выбирает в меню пункт «Подать статью» и вводит все необходимые данные. Своих соавторов при этом он выбирает из списка зарегистрированных пользователей.

Рукопись статьи готовится в соответствии с правилами оформления в редакторе MS Word и прикрепляется в формате *.doc, *.docx.

Сканированные лицензионный договор с подписями авторов и экспертное заключение (цветной режим сканирования, разрешение не менее 600 dpi) необходимо также разместить на сайте журнала в разделе «Подать статью» в формате *.pdf, *.jpg, *.jpeg.

По окончании всех работ обязательно нажать кнопку «Отправить в редакцию».

В редакцию журнала представляются следующие материалы.

1. **Статья**, подготовленная в соответствии с правилами оформления, – печатная версия, 2 экземпляра, подписанных авторами.

2. **Контактная информация** (телефоны рабочий и сотовый, адрес электронной почты, место работы, адрес места работы, должность, ученая степень, ученое звание автора) – печатная версия, 2 экземпляра.

3. **Описание статьи для базы данных «Российский индекс научного цитирования (РИНЦ)»**, подготовленное в соответствии с правилами оформления, – печатная версия, один экземпляр.

4. **Лицензионный договор**, заполненный и подписанный.

5. **Электронная версия статьи, контактной информации, описания статьи для базы данных РИНЦ, сканированный лицензионный договор и экспертное заключение о возможности опубликования** отправляются отдельными файлами на адрес редакции.

6. **Согласие на публикацию, обработку и распространение персональных данных авторов статей.**

7. **Экспертное заключение о возможности опубликования.**

Редакцией рассматриваются только те материалы авторов, которые полностью соответствуют вышеобозначенным требованиям. Неполный пакет материалов редакцией не рассматривается.

Подготовленные материалы направляются на почтовый адрес редакции: 630073, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет (НГТУ), корп. 7, ком. 606, в редакцию журнала «Безопасность цифровых технологий».

Все рукописи рецензируются, по результатам рецензирования редколлегия принимает решение о целесообразности опубликования материалов.

ВНИМАНИЕ!

Авторы несут ответственность за оформление, содержание и сам факт публикации статьи. Редакция журнала не несет ответственности за возможный ущерб, вызванный публикацией статьи. При наличии существенных недостатков в оформлении и содержании статьи редакция принимает решение об отклонении статьи без приведения полного перечня ошибок автора.

Ранее опубликованные материалы, а также материалы, представленные для публикации в других журналах, к рассмотрению не принимаются.

ПРАВИЛА ОФОРМЛЕНИЯ

При подготовке документов для отправки в редакцию журнала авторам рекомендуется внимательно прочитать правила и посмотреть примеры оформления статей и всех необходимых сопутствующих документов. Редакция рассматривает статьи, подготовленные как на русском, так и на английском языке. Для опубликования статьи на английском языке необходимо дополнительно предоставить ее русскоязычный вариант, оформленный по правилам журнала (кроме зарубежных авторов).

Перед отправкой рукописи в редакцию авторам необходимо проверить свою статью с помощью системы «Антиплагиат». Принятый редакционной коллегией уровень оригинальности статей должен составлять не менее 85 %.

Чтобы статья была направлена на рецензирование, необходимо подготовить следующее:

- 1) **статью** в соответствии с правилами оформления;
- 2) **контактную информацию** в одном файле предоставить по каждому автору: ФИО полностью, ученая степень, ученое звание, должность, место работы, адрес места работы, телефон рабочий и мобильный, адрес электронной почты;
- 3) **описание статьи для базы данных «Российский индекс научного цитирования (РИНЦ)»;**
- 4) **лицензионный договор** заполнить, бланк лицензионного договора должен быть подписан только авторами (он доступен авторам также в личном кабинете). Если авторов несколько, то необходимо добавить поля на всех авторов и подписать каждому из них;

5) **экспертное заключение** о возможности опубликования, принятое в вашей организации;

6) согласие на публикацию, обработку и распространение персональных данных авторов статей;

7) авторы, не являющиеся сотрудниками НГТУ, предоставляют **сопроводительное письмо** на имя проректора по научной работе НГТУ (ссылка на страницу сайта НГТУ). Письмо нужно подготовить на бланке организации с подписью и печатью руководителя.

ОСНОВНЫЕ РАЗДЕЛЫ ЖУРНАЛА

Автоматизация и управление технологическими процессами и производствами.

Управление в социальных и экономических системах.

Методы и системы защиты информации, информационная безопасность.

RULES FOR AUTHORS

CONDITIONS FOR ACCEPTANCE OF ARTICLES

All articles and their accompanying materials are submitted to the magazine through the magazine's website in electronic form after registration of all the authors of the article. Registration obliges each author to have an international ORCID. No other material supply options are considered.

The author (one of the co-authors) in his personal account selects the item "Submit article" in the menu and enters all the necessary data. At the same time, he selects his co-authors from the list of registered users.

The manuscript of the article is prepared in accordance with the design rules in the MS Word editor and attached in the format *.doc, *.docx.

Scanned license agreement with signatures of authors and expert opinion (color mode scanning, resolution not less than 600 dpi) can also be attached on the website of the magazine in the section "Submit article" in the format *.pdf, *.jpg, *.jpeg.

At the end of all works, be sure to click the "Send to Design" button.

The following materials are provided to the journal editor:

1. **The article**, prepared in accordance with the rules of design, is a private version, 2 copies signed by the authors.

2. **Contact information** (working and cellular phones, e-mail addresses, place of work, address of the place of work, position, scientific degree, academic title of the author) – printed version, 2 copies.

3. **The description of the article** for the database "**Russian Scientific Citation Index (RSCI)**", prepared in accordance with the rules of form-making, is a printed version, one copy.

4. **License agreement** completed and signed.

5. **Electronic version of the article**, contact information, description of the article for the RSCI database, scanned license agreement and expert opinion on the possibility of publication (in separate files to the editorial address).

6. **Consent to the publication, processing and dissemination of the personal data** of the authors of the articles.

7. **Expert opinion** on the possibility of publication.

The editors consider only those materials of the authors that fully meet the above requirements. Incomplete package of materials is not considered by the revision.

The prepared materials are sent to the postal address of the editorial office: 630073, Novosibirsk, Karl Marx Prospekt, 20, Novosibirsk State Technical University (NSTU), building 7, office 606, to the editors of the journal "Digital Technology Security".

All manuscripts were reviewed, and according to the results of the review, the editorial board decided on the appropriateness of publishing the materials.

ATTENTION!

The authors are responsible for the design, content and the fact of publication of the article. The editorial board of the journal is not responsible for possible damage caused by the publication of the article. If there are significant shortcomings in the design and content of the article, the editorial board decides to reject the article without giving a full list of the author's mistakes.

Previously published materials, as well as materials submitted for publication in other journals, are not accepted for consideration.

FORMATTING RULES

When preparing documents for submission to the journal editor, authors are advised to carefully read the rules and see examples of the design of articles and all necessary related documents. The Drafting Committee considered articles prepared in both Russian and English. To publish the article in English, it is necessary to additionally provide its Russian-language version, drawn up according to the rules of the magazine (except for foreign authors).

Before sending the manuscript to the editorial office, authors must check their article using the Antiplagiarism system. The level of originality of articles adopted by the Editorial Board should be at least 85 %.

For the article to be aimed at peer review, you need to prepare the following:

- 1) **the article** in accordance with the rules of design (volume from 7 to 30 pages);
- 2) **provide contact information** in one file for each author: full name, degree, academic title of the author, position, place of work, address of the place of work, telephone number of the worker and mobile, e-mail address;
- 3) **description of the article** for the database "Russian Scientific Citation Index (RSCI)";
- 4) fill out the **license agreement**, the form of the license agreement must be signed only by the authors (it is also available to the authors in the personal office), if there are several authors, then it is necessary to add fields on all authors and sign each of them;
- 5) **expert opinion** on the possibility of publication, adopted in your organization;
- 6) consent to the publication, processing and dissemination of the personal data of the authors of the articles;

7) authors who are not employees of the NSTU provide a **companion letter** addressed to the vice-rector for scientific work of the NSTU (link to the page of the NSTU website). The letter should be prepared on the form of the organization with the signature and seal of the manager.

JOURNAL SECTION

Automation and control of technological processes and productions.

Governance in social and economic systems.

Methods and systems of information protection, information security.