Научно-технический журнал

H&ES RESEARCH

Scientific and Technical Journal

НАУКОЕМКИЕ ТЕХНОЛОГИИ В КОСМИЧЕСКИХ ИССЛЕДОВАНИЯХ ЗЕМЛИ

HIGH TECHNOLOGIES IN EARTH SPACE RESEARCH

Журнал **H&ES Research** издается с 2009 года, освещает достижения и проблемы российских инфокоммуникаций, внедрение последних достижений отрасли в автоматизированных системах управления, развитие технологий в информационной безопасности, исследования космоса, развитие спутникового телевидения и навигации, исследование Арктики. Особое место в издании уделено результатам научных исследований молодых ученых в области создания новых средств и технологий космических исследований Земли.

Журнал H&ES Research входит в перечень изданий, публикации в которых учитываются Высшей аттестационной комиссиейРоссии (ВАК РФ), в систему российского индекса научного цитирования (РИНЦ), а также включен в Международный классификатор периодических изданий.

Тематика публикуемых статей в соответствии с перечнем группспециальностей научных работников по Номенклатуре специальностей:

- 2.2.15 Системы, сети и устройства телекоммуникаций (техн. науки)
- 2.3.1 Системный анализ, управление и обработка информации (техн. науки)
- 2.3.5 Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (техн. науки)
- 2.3.6 Методы и системы защиты информации, информационная безопасность (техн. науки)
- 2.5.13 Проектирование, конструкция и производство летательныхъ аппаратов (техн. науки)
- 2.5.16 Динамика, баллистика, управление движением летательных аппаратов (техн. науки)

ИНДЕКСИРОВАНИЕ ЖУРНАЛА H&ES RESEARCH

H&ES Research is published since 2009. The journal covers achievements and problems of the Russian infocommunication, introduction of the last achievements of branch in automated control systems, development of technologies in information security, space researches, development of satellite television and navigation, research of the Arctic. The special place in the edition is given to results of scientific researches of young scientists in the field of creation of new means and technologies of space researches of Earth.

The journal H&ES Research is included in the list of scientific publications, recommended Higher Attestation Commission Russian Ministry of Education for the publication of scientific works, which reflect the basic scientific content of candidate and doctoral theses. IF of the Russian Science Citation Index.

Subject of published articles according to the list of branches of science and groups of scientific specialties in accordance with the specialties:

- 2.2.15 Telecommunication systems, networks and devices
- 2.3.1 System analysis, management and information processing
- 2.3.5 Mathematical and software support for computing systems, complexes and computer networks
- 2.3.6 Methods and systems of information security
- 2.5.13 Design, construction and production of aircraft
- 2.5.16 Dynamics, ballistics, aircraft motion control

JOURNAL H&ES RESEARCH INDEXING

• NEICON • CyberLenika (Open Science) • Google Scholar • OCLC WorldCat • Ulrich's Periodicals Directory • • Bielefeld Academic Search Engine (BASE) • eLIBRARY.RU • Registry of Open Access Repositories (ROAR)

Все номера журнала находятся в свободном доступе на сайте журнала **www.hes.ru** и библиотеке **elibrary.ru.**

Всем авторам, желающим разместить научную статью в журнале, необходимо оформить ее согласно требованиям и направить материалы на электронную почту: **HT-ESResearch@yandex.ru.**
С требованиями можно ознакомиться на сайте: **www.H-ES.ru.**

Язык публикаций: русский, английский.
Периодичность выхода – 6 номеров в год.
Свидетельство о регистрации СМИ ПИ № ФС 77-60899 от 02.03.2015
Территория распространения: Российская Федерация, зарубежные страны

Тираж 1000 экз. Цена 1000 руб.
Плата с аспирантов за публикацию рукописи не взимается.

All issues of the journal are in a free access on a site of the journal **www.hes.ru** and **elibrary.ru.**

All authors wishing to post a scientific article in the journal, you must register it according to the requirements and send the materials to your email: **HT-ESResearch@yandex.ru.**
The requirements are available on the website: **www.H-ES.ru.**

Language of publications: Russian, English.
Periodicity – 6 issues per year.
Media Registration Certificate PI No. FS77-60899. Date of issue: March 2, 2015.
Distribution Territory: Russian Federation, foreign countries

Circulation of 1000 copies. Price of 1000 Rur.
Postgraduate students for publication
of the manuscript will not be charged

ISSN 2409-5419 (Print) ISSN 2412-1363 (Online) WWW.H-ES.RU Дата выхода: 20 июня 2023

T.15
4
2023

Учредитель:
ООО "ИД Медиа Паблишер"

Издатель:
ДЫМКОВА С.С.

Главный редактор:
ЛЕГКОВ К.Е.

Адрес издателя:
111024, Россия, Москва,
ул. Авиамоторная, д. 8, корп. 1, офис 323.

Адрес редакции:
194044, Россия, Санкт-Петербург,
Лесной Проспект, 34-36, к. 1,
Тел.: +7(911) 194-12-42.

Адрес типографии:
Россия, Москва, ул. Складочная, д. 3, кор. 6.

СОДЕРЖАНИЕ

CONTENTS

Founder:
"Media Publisher", LLC

Publisher:
DYMKOVA S.S.

Address of publisher:
111024, Russia, Moscow,
st. Aviamotornaya, 8, bild. 1, office 323

Address of edition:
194044, Russia, St. Petersburg,
Lesnoy av., 34-36, h.1,
Phone: +7 (911) 194-12-42.

Address of printing house:
Russia, Moscow, st. Skladochnaya, 3, h. 6

doi: 10.36724/2409-5419-2023-15-4-4-13

АВТОМАТИЧЕСКОЕ ПОСТРОЕНИЕ ОТКАЗОУСТОЙЧИВЫХ БОРТОВЫХ СЕТЕЙ

ОЛЕНЕВ
Валентин Леонидович[1]

АННОТАЦИЯ

Введение: отказоустойчивость является одним из ключевых вопросов для бортовых сетей автономно функционирующих аппаратов. С ростом количества устройств в составе таких сетей появляется потребность в методах автоматизации построения сетей с учетом необходимой отказоустойчивости и ограничений. **Цель исследования:** целью исследования является реализация методов и алгоритмов, способствующих оценке отказоустойчивости бортовых сетей, а также, при необходимости, автоматического достраивания сетей до требуемой отказоустойчивости. **Методы:** полученные методы и алгоритмы основаны на элементах теории графов, а именно оценке связности и поиска кратчайших путей. Графы составляются на основе структуры бортовой сети с применением ряда предложенных правил. В качестве метода обеспечения отказоустойчивости рассматривается структурная избыточность, которая при помощи разработанного метода сводится к обоснованному дополнению сети элементами, в отличие от часто применяемого полного копирования сетевой структуры. Метод состоит из двух этапов, в рамках которых производится достраивание сети до требуемой отказоустойчивости, а затем итерационная доработка структуры сети с целью уменьшения аппаратных затрат. **Результаты:** использование разработанного метода позволяет в значительной степени упростить и ускорить проектирование бортовых сетей, поскольку формализованная оценка такой сети проводится еще на этапе проектирования, до реальной ее сборки. Процесс достраивания до необходимой отказоустойчивости для крупных сетевых структур занимает меньше минуты. **Практическая значимость:** представленный метод реализован в рамках программного комплекса автоматизированного проектирования и моделирования бортовых сетей, произведена длительная апробация в составе реальных проектов по разработке бортовых сетей космических аппаратов. **Обсуждение:** аналогов подобной программной реализации для бортовых сетей не существует. Полученные при помощи метода решения на практике показали себя близкими к оптимальным и получили высокую оценку специалистов.

Сведения об авторе:

[1] к.т.н., доцент, заведующий кафедрой аэрокосмических компьютерных и программных систем, директор центра аэрокосмических исследований и разработок, Федеральное государственное автономное образовательное учреждение высшего образования "Санкт-Петербургский государственный университет аэрокосмического приборостроения", г. Санкт-Петербург, Россия, Valentin.Olenev@guap.ru

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, соглашение № FSRF-2023-0003, "Фундаментальные основы построения помехозащищенных систем космической и спутниковой связи, относительной навигации, технического зрения и аэрокосмического мониторинга".

КЛЮЧЕВЫЕ СЛОВА: *отказоустойчивость, структурная избыточность, бортовые сети, автоматизация, достраивание сети.*

Для цитирования: *Оленев В.Л.* Автоматическое построение отказоустойчивых бортовых сетей // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 4-13. doi: 10.36724/2409-5419-2023-15-4-4-13

Введение

Создание отказоустойчивых сетей является одной из ключевых проблем в тех случаях, когда работа сети критически важна для функционирования объекта, в работе которого она задействована. В качестве таких объектов могут выступать крупные предприятия, атомные станции, бортовые сети летательных и космических аппаратов, автономных роботов и автомобилей и т.п. [1, 2, 3]. При этом возникновение отказов допускается, полностью избежать их невозможно, особенно при длительном сроке функционирования. Но необходимы эффективные методы парирования этих отказов и устранения их последствий.

Вопросы, связанные с отказоустойчивостью локальных вычислительных сетей уже в достаточной степени изучены, но для бортовых сетей, использующих специализированные коммуникационные протоколы, имеющих значительные ограничения по массе и энергопотреблению, эти вопросы остаются крайне актуальными.

Отказоустойчивость – это свойство архитектуры информационных систем, обеспечивающее выполнение заданных функций в случаях, когда в аппаратных и программных средствах системы возникают отказы [4]. Отказоустойчивость может быть активная или пассивная. Активная отказоустойчивость подразумевает наличие специализированных средств обнаружения, локализации отказа с последующей реконфигурацией системы, чаще в автоматическом режиме, поэтому является ресурсозатратной. Пассивная отказоустойчивость заключается в дополнении системы резервной аппаратурой с возможностью максимально быстрого переключения между ними.

Введение дополнительного оборудования называется избыточностью, наиболее эффективным ее видом является структурная избыточность. Этот подход дает больший вес сети благодаря дополнительному оборудованию, но является более надежным и повсеместно применяется в бортовых сетях, где количество резервной аппаратуры иногда превышает количество основной. Таким образом, любая система, содержащая избыточные компоненты или функции обладает некоторыми свойствами отказоустойчивости [5].

Отказоустойчивость в бортовых сетях

С каждым новым этапом развития бортовых сетей и использующейся в них техники количество узлов, включенных в работу сети, растёт. Узлы могут представлять из себя различную по сложности и степени интеллектуальности аппаратуру, например, процессоры, датчики, блоки управления, память и т.д. Тенденции развития идут к тому, что все эти устройства будут объединены в одну сеть, по которой будет проводиться обмен данными между всеми этими устройствами. Потеря данных в бортовой сети или отказ одного из этих устройств может повлечь серьезные последствия, поэтому в таких сетях необходимо вводить дополнительные каналы связи, обеспечивающие резервные маршруты, а также дополнять топологию сети резервными устройствами [6].

Глобально бортовая сеть состоит из трех основных элементов: терминальные узлы, коммутаторы и каналы.

Каждый из этих представленных элементов может отказать без возможности его восстановления в разумные временные сроки. Продолжение корректного функционирования при отказах и сбоях достигается избыточностью. Например, один терминальный узел может быть представлен двумя или тремя комплектами аппаратуры, дублирующими друг друга по функциональности [7]. При отказе одного из комплектов на замену ему включается резервный комплект, который продолжает выполнять необходимые задачи.

В бортовых системах широко применяются дискретные детерминированные показатели отказоустойчивости, отражающие, какое число отказов компонентов может выдержать проектируемая система, сохраняя работоспособное состояние. Для этого используется показатель *d*-безотказности, который показывает, что при любых *d* отказах внутри системы она корректно функционирует, но найдется такое сочетание *d*+1 отказов, при которых её алгоритм нарушается [8].

Величину *d* и величину вероятности безотказной работы в течение определенного промежутка времени связывает некоторая (особая для каждого метода) функциональная зависимость. Ряд решений, направленных на введение избыточности в бортовые сети, был предложен в работах [9, 10, 11]. Однако, эти решения не дают однозначного формализованного алгоритма действий или метода для автоматизированного получения сети, устойчивой к *d* отказам. При этом задача автоматизации процесса построения отказоустойчивых бортовых сетей становится всё более актуальной с ростом количества автономно функционирующих средств и сокращающихся сроков их реализации.

Таким образом, целью данного исследования является разработка метода автоматизации процесса построения отказоустойчивых сетей. Применительно к решаемой задаче будут рассматриваться отказы коммутаторов и каналов связи, поскольку отказ устройства или комплекта может быть приравнен к отказу связанного с этим комплектом канала с точки зрения структуры сети. При этом в рамках данной задачи не рассматривается тип резерва, в котором находятся устройства (холодный или горячий резерв), поскольку он также не оказывает влияния на сетевую структуру. В соответствии с требованиями индустрии, изложенными в [12], на текущий момент количество комплектов на устройстве может быть минимум 1, максимум 3. Пример отказоустойчивой связи между двумя сетевыми узлами, состоящими из двух комплектов через коммутатор показан на рисунке 1. При этом каждый из комплектов имеет основной (О) и резервный (Р) порты.

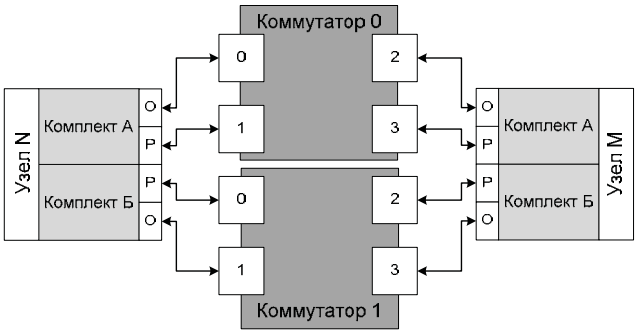


Рис. 1. Пример отказоустойчивой связи

Представление сети в виде графа

Проблема анализа отказоустойчивости сетей связана с вычислением связности топологии. Такой формализованный подход используется в большом количестве научных и инженерных задач [13]. Этот же подход лёг в основу данного исследования.

Для начала, необходимо представить сеть виде графа $G = (V, E)$, где V – это множество узлов и коммутаторов в сети, а E – множество каналов связи. При этом должно учитываться количество комплектов на каждом устройстве. Начальная топология сети задается проектировщиком. Он составляет произвольную структуру, которая может быть представлена как двумя узлами, соединенными каналами связи, так и распределенной сетевой структурой, использующей некоторое количество коммутаторов.

Пусть f – это количество отказов, к которым устойчива сеть. Задача оценки отказоустойчивости сети сводится к определению k-связности графа [4**Ошибка! Закладка не определена.**], который строится по исходной топологии сети. Отказоустойчивость вычисляется в соответствии с формулой 1.

$$f = k - 1, \tag{1}$$

где f – количество отказов; k – связность графа.

Чем больше связность, тем большее количество элементов сети может отказать, и это не окажет негативного влияния на функционирование сети [14]. В графе должно быть более одного маршрута между каждой парой вершин для парирования отказов, это значит, что граф должен оставаться связным при удалении ребра или вершины.

Для графов определено два вида связности: вершинная и реберная. Вершинная связность графа есть минимальное число вершин, которые нужно удалить, чтобы разделить этот граф на большее количество компонент связности (минимум две части). Реберная связность – минимальное число ребер, которые нужно удалить с этой же целью [15]. Отказ канала можно считать равнозначным отказу порта на терминальном узле или коммутаторе. Может отказать узел, его комплект, порт комплекта или коммутатор. Следовательно, отказоустойчивость бортовой сети требует определения вершинной связности графа.

Для случая двухточечного соединения (подключения двух узлов типа точка-точка, *point-to-point*) [16] оценка отказоустойчивости выполняется по Правилу 1.

Правило 1: Количество отказов, к которому устойчива система точка-точка, вычисляется исходя из минимального количества комплектов на двух устройствах в системе, связанных с соседним устройством каналами связи (U).

В этом случае количество отказов вычисляется по формуле 2.

$$f = U - 1, \tag{2}$$

где f – количество отказов; U – минимальное количество комплектов на двух устройствах.

Пример для случая точка-точка приведен на рисунке 2.

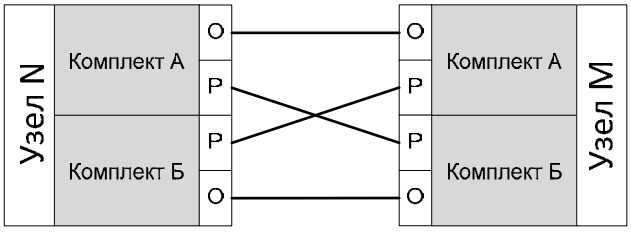


Рис. 2. Пример связи для случая точка-точка

Система на рисунке 2 представлена двумя узлами, содержащими по два комплекта. Каждый из комплектов имеет по два порта, все порты двух узлов связаны каналами. Производя расчет по формуле 2, получаем, что сеть устойчива к одному отказу (может отказать комплект целиком).

Рассмотрим процесс формирования графа на основе топологии сети, где сеть содержит коммутаторы. Граф формируется в соответствии с Правилом 2, результат применения которого приведен на рисунке 3.

Правило 2:

1. Каждый узел должен быть представлен одной вершиной в графе.
2. При наличии в узле нескольких комплектов они считаются как один узел.
3. Каждый коммутатор должен быть представлен отдельной вершиной в графе.
4. Каждый канал, соединяющий элементы сети, должен быть поставлен в соответствие ребру графа, соединяющему соответствующие вершины графа.

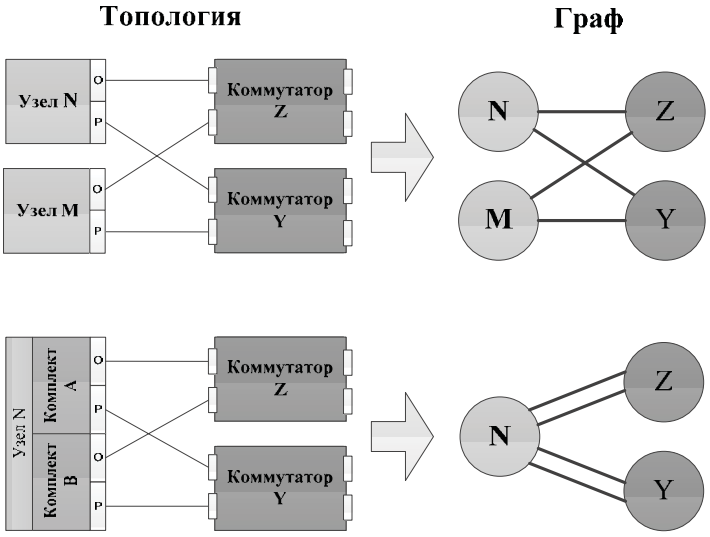


Рис. 3. Переход от топологии к графу

Связность анализируется на ориентированных графах [17]. Современные стандарты, предназначенные для надежной передачи данных, имеют двунаправленные каналы связи (для отправки сервисных данных). С учетом этого, каждый канал представляется парой противоположно направленных ребер.

Пример показан на рисунке 4.

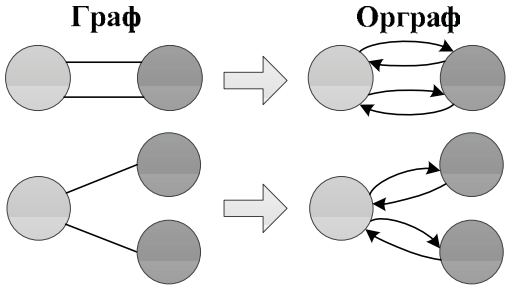


Рис. 4. Переход к ориентированному графу

Метод оценки отказоустойчивости коммуникационной сети

После построения графа сетевой топологии производится выполнение оценки отказоустойчивости. Метод состоит из нескольких шагов:

1. Проверка графа на связность. При несвязном графе сеть не устойчива к ошибкам. Если граф связен, переход на шаг 2.

2. Вычисление вершинной связности графа. Для этого необходимо перебрать все пары вершин, найти количество вершинно-непересекающихся путей и выбрать минимум, который соответствует значению связности графа *k*. Алгоритм вычисления вершинной связности представлен в [18].

3. Подсчет отказоустойчивости. Вычисляется количество отказов, к которым устойчива сеть, в соответствии с формулой 1.

Приведенный метод решает конкретную задачу автоматизации определения отказоустойчивости сети. Однако, возможна ситуация, при которой полученная отказоустойчивость будет равна 0. Это значит, что в сети есть узкие места, которыми могут быть как коммутаторы, так и каналы.

Каналы сети, которые являются ее узким местом, соответствуют мостам графа, то есть ребрам, после удаления которых граф распадается на два не связанных подграфа. Поиск мостов выполняется с помощью алгоритма поиска в глубину [19]. Ребро будет являться мостом тогда и только тогда, когда оно присутствует в дереве обхода в глубину и из вершины и любого ее потомка нет обратного ребра в вершину или ее предка. Это значит, что при удалении ребра между таким узлом и его предком будет отсоединено поддерево. Обнаруженные мосты – это каналы, являющиеся узкими местами в сети.

Отказ коммутатора также может привести к разделению сети на две подсети. Такие коммутаторы соответствуют точкам сочленения в графе, то есть вершинам, в случае удаления которой связный граф распадается на два (или больше) непересекающихся подграфа [20]. Поиск точек сочленения также выполняется на основе алгоритма поиска в глубину. Обнаруженные точки сочленения – это коммутаторы, которые являются «узкими местами» в топологии.

Приведенный метод оценки отказоустойчивости коммуникационной сети дает возможность автоматизировать процесс определения, сколько отказов выдержит проектируемая бортовая сеть. Однако, во многих случаях проектировщик самостоятельно не сможет преобразовать или исправить сеть, если она не отвечает его требованиям отказоустойчивости. Для этого необходима разработка дополнительного формализованного подхода.

Метод достраивания структуры сети до необходимой отказоустойчивости

В рамках предложенного метода выполняется автоматизированная трансформация сети для обеспечения требуемого уровня отказоустойчивости. Задача трансформации сети для обеспечения отказоустойчивости формулируется следующим образом: необходимо добавить избыточные комплекты для терминальных узлов, разместить дополнительные коммутаторы и коммуникационные каналы в исходную топологию сети для достижения требуемой отказоустойчивости. Могут быть определены отдельные участки сети, которые не участвуют в трансформации для обеспечения отказоустойчивости.

Данная задача будет решаться в два этапа:

1) формирование топологии сети, устойчивой к заданному количеству отказов;

2) итерационная доработка полученной топологии сети.

Каждый из этих этапов методики является достаточно громоздким, поэтому рассмотрим эти этапы по отдельности.

Для формального описания разработанной методики примем следующие обозначения:

$N = (N_1, N_2, ..., N_l)$ – множество всех узлов в сети (или в ее части).

$R = (R_1, R_2, ..., R_m)$ – множество всех коммутаторов в сети (или в ее части).

$Ch = (Ch_1, Ch_2, ..., Ch_s)$ – множество всех каналов в сети (или в ее части).

$Net = N \cup R \cup Ch$ - сеть состоит из множества узлов, коммутаторов и каналов

np – количество недостающих портов в узле для обеспечения требуемой отказоустойчивости.

F – требуемая отказоустойчивость;

F_c – текущая отказоустойчивость;

cp – текущее количество портов в узле;

FreePorts () – алгоритм, который вычисляет количество несвязанных портов на устройстве;

R_{pi} – множество коммутаторов-соседей i-го порядка;

(P_{main}, R_{red}) – порт соединен (*true*) или отсоединен (*false*) от коммутатора;

N_fports – количество свободных портов у узлов;

R_fports – количество свободных портов у коммутаторов;

$port_d$(R_m или N_l) – функция определения, подключен порт или нет, возвращает *true*/*false*.

Этап 1. Формирование сети, устойчивой к заданному количеству отказов

Для описания метода автоматизированного достраивания сети до заданной отказоустойчивости необходимо разработать ряд алгоритмов, решающих промежуточные задачи.

К таким алгоритмам относятся:

1) Алгоритм проверки необходимого количества комплектов в узлах сети.

2) Алгоритм автоматического добавления каналов связей между коммутаторами.

3) Алгоритм автоматического достраивания коммутаторов в сети.

Рассмотрим каждый из этих алгоритмов отдельно.

Алгоритм проверки необходимого количества комплектов в узлах сети

Для проверки необходимого количества комплектов в узлах сети выполняются следующие действия. Для всех узлов в составе сети происходит последовательная проверка подключения основного порта основного комплекта к коммутатору, находящемуся в холодном резерве. Если это так, то пользователю должно быть выведено предупреждение об этом. Далее выполняется проверка необходимости добавления комплектов в узел. Вычисление количества недостающих портов выполняется по формуле 3.

$$np = (F+1) - cp\,, \quad (3)$$

где np – количество недостающих портов в узле; F – требуемая отказоустойчивость; cp – текущее количество портов в узле.

В зависимости от полученного значения np будет добавлено различное количество комплектов:

• Если np меньше 1, то нет необходимости в добавлении комплектов. Далее проверяется, что порты комплектов в составе узлов подключены к разным коммутаторам. Если это не так, то необходимо подключить резервный порт комплекта к соседнему коммутатору первого порядка, в случае невозможности – к соседу второго порядка и т.д. Количество коммутаторов, к которым подключен узел, должно быть больше, чем требуемая отказоустойчивость, в противном случае необходимо отключить все порты узла, кроме основного порта основного комплекта, и попытаться подключить их к другим коммутаторам.

• Если np больше 1, то для данного узла необходимо создать дополнительные комплекты с теми же свойствами, что и у основного. Для $1 \le np \le 2$ создается один дополнительный комплект, для $np \ge 3$ – два комплекта. Всегда создаются комплекты с двумя портами. Если устройство имело один порт, то после добавления комплектов, оно будет содержать все комплекты с двумя портами.

После этого выполняется подключение новых комплектов к коммутаторам. Если у коммутаторов свободных портов достаточно, то выполняется связывание, иначе – выход. Если подключить к соседнему коммутатору не удается, то производится подключение к соседу второго порядка и т.д.

Формальная запись алгоритма проверки необходимого количества комплектов в узлах сети (*CheckUnits*) выглядит следующим образом.

Входные данные: *Net, F*; выходные данные: *Net'*.

Алгоритм 1: Проверка необходимого количества комплектов в узлах сети

0: **for** $\forall N \in Net : (P_{main}, R_{main}), P_{main} \in U_{main}$ **do**
1: **if** $(P_{main}, R_{red}) = true$ **then**
2: *Print*(«Основной порт основного комплекта узла N подключен к коммутатору *R*, находящемуся в холодном резерве»)
3: **endif**
4: **endfor**
5: $np \leftarrow (F+1) - cp$
6: **If** $np < 1$ **then**
3: **If** $\exists N \in Net : (p_{main}, R) = true \wedge (p_{res}, R) = true$ **then**
4: $i \leftarrow 1$
4: $(p_{res}, R) = false$
5: **while** $FreePorts(R_{pi}) = 0$
6: $R' \in R_{pi}$
7: $i \leftarrow i+1$
8: **endwhile**
9: **endif**
10: $(p_{res}, R'); R' \neq R; R' \in R_{pi}$
11: **endif**
12: **If** $|R| \le F$ **then**
13: $((P,R) = false) \wedge ((p_{main}, R) = true); (P, R_x) = true; R \neq R_x$
14: **else** $AddUnits(N, np-1)$
15: **endif**
16: $N\_fports \leftarrow FreePorts(N)$
17: $R\_fports \leftarrow FreePorts(R)$
18: **If** $R\_fports < N\_fports$ **then**
19: $AddChannels(Net, F)$
20: **else** $i \leftarrow 1$
21: **while** $FreePorts(R_{pi}) = 0$
22: $R'' \in R_{pi}$
23: $i \leftarrow i+1$
endwhile
24: $\forall N \in Net : (p_{new}, R'') = true, R'' \in R_{pi}$
25: **endif**

Алгоритм автоматического добавления каналов связей между коммутаторами

Если у коммутаторов в обрабатываемой сети есть свободные порты, то необходимо связать каналами все коммутаторы, еще не связанные друг с другом (по одному каналу связи на каждую пару).

Формальная запись алгоритма автоматического добавления каналов связей между коммутаторами (*AddChannels*) выглядит следующим образом.

Входные данные: *Net*; выходные данные: *Net'*.

Алгоритм 2: Автоматическое добавление каналов связей между коммутаторами

0: **If** $\forall R_i \in Net : FreePorts(R_i) > 0$ **then**
1: **If** $\forall R_j \in Net : (FreePorts(R_j) > 0) \wedge ((R_i, R_j) = false)$ **then**
2: $(R_i, R_j) = true$
3: **endif**
4: **endif**

Добавление каналов между коммутаторами может повысить связность сети, а как следствие – отказоустойчивость. У части коммутаторов могут отсутствовать свободные порты, поэтому для них добавление новых каналов невозможно.

Алгоритм автоматического достраивания коммутаторов в сети

В первую очередь выполняется проверка наличия неподключенных коммутаторов, которые могли появиться при отключении портов. Если таковые найдены, то необходимо произвести их подключение к основной части сети.

Добавление коммутаторов происходит посредством создания копий сетевой структуры вместе с каналами между коммутаторами. Количество необходимых дополнительных копий вычисляется как разность требуемой и текущей отказоустойчивости. Копируя структуру коммутаторов, создаются копии для каждого из них, а также аналогичные каналы связи, что и в исходной сети. При этом не копируются узлы и связи с узлами. После этого копии коммутаторов связываются друг с другом при наличии свободных портов.

Если у узлов есть неподключенные порты, то выполняется связывание узлов и коммутаторов при наличии свободных портов для того, чтобы подключить все узлы. Если свободных портов достаточно, то выполняется связывание, иначе проектировщику сети выводится сообщение, что достроить сеть не удается и необходимо использовать коммутаторы с большим количеством портов.

Формальная запись алгоритма автоматического достраивания коммутаторов (*AddRouters*) выглядит следующим образом.

Входные данные: *Net*, *F*; выходные данные: *Net'*.

Алгоритм 3: Автоматическое достраивание коммутаторов в сети

0: **If** $\forall R_i \in Net : ((R_i, p_j) = false) \wedge ((R_i, R_k) = false)$ **then**

1: $(R_i, R_k) = true$

2: **endif**

3: $i \leftarrow 0$

4: **while** $i < (F - F_c)$

5: $Net' \leftarrow Net \cup R_k; R_k = R_i; R_i \in Net$

6: $i \leftarrow i + 1$

7: **endwhile**

8: $\forall R_k, R_t \in Net' : (R_t, R_k) = true$

9: $N\_fports \leftarrow FreePorts(N)$

10: $R\_fports \leftarrow FreePorts(R)$

11: **If** $R\_fports < N\_fports$ **then**

12: $AddChannels(Net, F)$

13: **else** Print("С данным количеством узлов и портов у коммутаторов достроить сеть не удается. Используйте коммутаторы с большим количеством портов")

14: **endif**

15: **If** $\forall N_i \in Net : FreePorts(N_i) <> 0$ **then**

16: **If** $((N_i, R_k) = true) \wedge (R_k \neq R_{k+1})$ **then**

17: $(N_i, R_{k+1}) = true$

18: **endif**

19: **endif**

Достраивание сети до необходимой отказоустойчивости

Разработанный метод достраивания структуры сети до необходимой отказоустойчивости использует вышеописанные алгоритмы и состоит из основных четырех шагов:

1) Оценка сети на отказоустойчивость. Если отказоустойчивость меньше, чем требуемая – перейти на следующий шаг, иначе закончить работу.

2) Проверка необходимого количества комплектов у узлов для обеспечения требуемой отказоустойчивости при помощи разработанного алгоритма. Если все комплекты всех узлов подключены к коммутаторам, то проверяется отказоустойчивость сети. Если отказоустойчивость сети соответствует требуемой, то задача выполнена, в противном случае, а также, если остались неподключенные порты, выполнить переход на шаг 3.

3) Проверить количество коммутаторов в сети.

a. Если количество коммутаторов в сети больше требуемой отказоустойчивости:

o Если все комплекты узлов подключены к коммутаторам, то в случае, если у коммутаторов есть свободные порты, то используется алгоритм добавления каналов связей между коммутаторами. Если добавлены новые каналы между коммутаторами, то проверяется отказоустойчивость. Если она стала соответствовать требуемой, то выполняется переход на Этап 2. Иначе необходимо добавлять дополнительные коммутаторы в сеть, отключив все порты узлов, которые были связаны с коммутаторами на шаге 2. Все исходные связи необходимо оставить прежними.

o Если у коммутаторов нет свободных портов, то выполняется алгоритм достраивания коммутаторов.

o У узлов есть порты, не подключенные к коммутаторам. В этом случае необходимо выполнить алгоритм достраивания коммутаторов.

b. Если количество коммутаторов в сети меньше или равна требуемой отказоустойчивости, то для дальнейшей трансформации сети необходимо выполнить алгоритм достраивания коммутаторов. В результате все порты узлов должны быть подключены к коммутаторам.

4) Проверка отказоустойчивости результирующей сети. Если отказоустойчивость равна или превышает требуемую, то необходимо выполнить переход на Этап 2. Если отказоустойчивость меньше требуемой, то проектировщику выводится сообщение о невозможность выполнить достраивание сети до требуемой отказоустойчивости.

Формально метод генерации отказоустойчивой структуры можно представить следующим образом:

1. $F_c = FT(Net)$

Если $F_c \geq F$ – окончание работы метода, иначе – п.2

2. Если $N \neq \varnothing, N \in Net$, то $Net' = CheckUnits(Net)$, иначе – выход

Если $\forall Ch_s = (U_j, R_m)$, то $F_c = FT(Net')$

Если $F_c \geq F$ – окончание работы метода, иначе – п.3

3. Если $m > F$

a. $\forall Ch_s = (U_j, R_m)$

- $\forall R_m$: Если $\exists port_d(R_m) = false$, то $Net'' = AddChannel(Net', F)$

$F_c = FT(Net'')$

Если $F_c \geq F$ – $Net \leftarrow Net''$,

иначе отмена $AddChannel(Net', F)$,

- Если $\forall port_d(R_m) = true$, $Net''' = AddRouters(Net'', F)$

b. Если $\exists port_d(N_l) = false$, то $Net''' = AddRouters(Net'', F)$

Если $m \leq F$

$Net''' = AddRouters(Net'', F)$

4. $F_c = FT(Net''')$.

Если $F_c \geq F$, то $Net \leftarrow Net'''$,

иначе *Print* («Невозможно выполнить достраивание сети до требуемой отказоустойчивости»).

Этап 2. Итерационная доработка топологии сети.

В процессе копирования структуры сети, ее каналов, коммутаторов, а также дальнейшем связывании элементов сети может возникнуть избыточность. Появятся элементы сети, необходимы в соответствии с вышеописанным методом, но при исключении из структуры сети которых можно получить сеть, более приближенную к оптимальной с точки зрения аппаратных затрат. Необходимо разработать вспомогательный метод итерационной доработки сети, который будет состоять в последовательном удалении добавленных элементов и последующей проверке отказоустойчивости.

Метод итерационной доработки топологии сети приведен ниже:

1) Сформировать граф на основе заданной топологии сети по Правилу 2.

2) Сформировать множество вершин графа, которые соответствуют вновь добавленным коммутаторам.

3) Сформировать множество ребер графа, которые соответствуют вновь добавленным каналам связей.

4) Если добавлены коммутаторы, то последовательно для каждого из них выполнить следующие действия:

a. Удалить текущий коммутатор из графа и все инцидентные ему ребра.

b. Выполнить оценку отказоустойчивости получившегося графа.

o Если удаление коммутатора приводит к удалению связей с одним или более узлов, то вернуть текущий коммутатор в граф вместе с удаленными ребрами.

o Если текущая отказоустойчивость больше или равна требуемой, то данный коммутатор должен быть удален из структуры сети вместе со всеми инцидентными ребрами.

o Если текущая отказоустойчивость меньше требуемой, то вернуть текущий коммутатор в граф вместе с удаленными ребрами.

c. Остановить данную проверку, когда будут рассмотрены все вновь добавленные коммутаторы.

5) Если были добавлены каналы, то последовательно для каждого из них выполнить следующие действия:

a. Удалить текущий канал, если это не приводит к удалению связи с узлом. При удалении необходимо удалить оба ребра орграфа, соответствующие данному каналу.

b. Выполнить оценку отказоустойчивости получившегося графа.

o Если текущая отказоустойчивость больше или равна требуемой, то канал должен быть удален.

o Если текущая отказоустойчивость меньше требуемой, то вернуть текущий канал в граф.

c. Остановить проверку, когда будут рассмотрены все вновь добавленные каналы.

6) Сохранить результирующую сеть

Формально метод итерационной доработки топологии сети выглядит следующим образом:

1. Сформировать граф *Net*
2. $R_{new} \in Net$
3. $Ch_{new} \in Net$
4. Для всех R_{new}:

Если $R_{new} \neq \varnothing$, то $F_c = FT(Net')$

$\forall R_i \in R_{new} : Net' := (R_i \cap Net) \wedge ((Ch_j(R_i)..Ch_n(R_i)) \cap Net)$

Если $F_c \geq F$, то

$\forall R_i \in R_{new} : Net' := (R_i \cap Net) \wedge ((Ch_j(R_i)..Ch_n(R_i)) \cap Net)$

иначе $Net \leftarrow Net'$

5. Для всех Ch_{new}:

Если $Ch_{new} \neq \varnothing$, то

$\forall Ch_i \in Ch_{new} : Net' := (Ch_i \cap Net)$

Если $N(Net') \neq N(Net)$, то $Net' \leftarrow Net$

$F_c = FT(Net')$

Если $F_c \geq F$, то $\forall Ch_i \in Ch_{new} : Net' := (Ch_i \cap Net)$,

иначе $Net' \leftarrow Net$;

6. $Net \leftarrow Net'$.

Результатом работы данного метода могут быть различные сетевые структуры. Чтобы принять решение, какой вариант сети более подходит к текущей задаче, проектировщик может оценить физические характеристики (количество коммутаторов, терминальных узлов, показатели массы и энергопотребления).

Применение метода ограничено в следующих случаях:

1) На каком-либо из узлов основной порт основного комплекта узла подключен к коммутатору, находящемуся в резерве.

2) Используются коммутаторы с недостаточным количеством портов, поэтому при добавлении новых коммутаторов и каналов все порты автоматически будут заняты дублирующими связями – повысить связность сети не удастся.

3) Другие частные случаи, при которых добавление резервных коммутаторов не увеличивает отказоустойчивость сети.

4) Задано требование по отказоустойчивости, не удовлетворяющее ограничениям.

Корректность разработанных методов следует из применения математического аппарата теории графов для исследования и разработки и результатов моделирования в составе реализованного программного обеспечения.

Экспериментальное подтверждение работоспособности разработанного метода

Предложенные методы, предназначенные для автоматизации процесса построения отказоустойчивых структур, были реализованы в рамках создания программного комплекса SANDS (*SpaceWire Automated Design and Simulation*) [21] по заказу АО «ИСС» им. академика М.Ф. Решетнева. Это автоматизированная система проектирования и моделирования сетей SpaceWire. Первый программный компонент системы включает в себя функциональность анализа сети на узкие места и отказоустойчивость, а также автоматическое достраивание сети до необходимой отказоустойчивости. На рисунке 6 показан пример сети, состоящей из трех узлов (по два комплекта), и двух коммутаторов (основной и резервный).

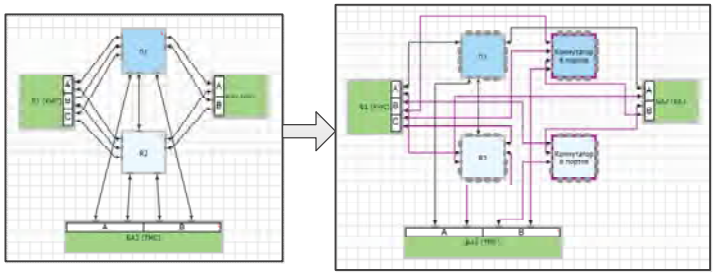


Рис. 5. Автоматизированное достраивание сети до необходимой отказоустойчивости

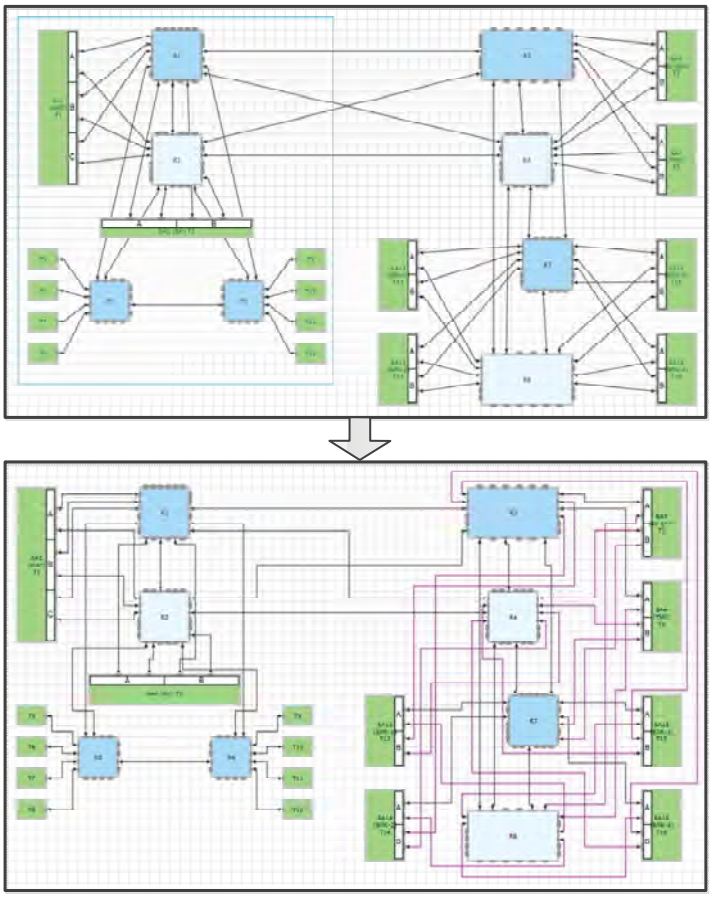


Рис. 6. Достраивание до необходимой отказоустойчивости участка сети

При запуске автоматизированной генерации с требованием достроить сеть до отказоустойчивости 2 добавлены каналы и два новых коммутатора, показанные фиолетовым цветом. Метод также может применяться только для отдельных областей сети. На рисунке 6 показан выбор области сети, не участвующей в оценке отказоустойчивости и достраивании. Остальная часть сети должна быть достроена до отказоустойчивости 3.

Таким образом, программная реализация разработанного метода подтверждает его работоспособность. Проведенное тестирование показывает, что большинство полученных решений получаются или оптимальными, или близкими к оптимальным (в зависимости от размера сети).

Заключение

Разработанный метод позволяет проводить оценку отказоустойчивости заданной сетевой структуры, которая может быть выполнена для сети в целом или для отдельных частей сети, определенных проектировщиком. Метод позволяет определять «узкие места» сети – коммутаторы и каналы связи, при выходе из строя которых сеть перестает быть устойчивой к отказам.

Также метод позволяет автоматически трансформировать сеть, заданную проектировщиком, для обеспечения требуемого уровня отказоустойчивости. В результате выполнения трансформации сети пользователю выдается новая структура сети, в которой могут быть добавлены комплекты на узлах, коммутаторы и каналы.

Метод достраивания структуры сети до необходимой отказоустойчивости отличается от уже существующих тем, что направлен на получения решений, близких к оптимальным по затратам аппаратуры, а также тем, что он разработан с учетом технических требований индустрии.

Реализация и успешная работа метода в рамках программного комплекса SANDS, подтверждает его состоятельность, метод выполняет поставленные задачи по достраиванию сетевых структур, близких по структуре к реальным бортовым сетям.

Подобные решения для автоматизации сложных процессов построения бортовых сетей становятся неотъемлемой частью технологии и процесса создания аппаратов, и в дальнейшем позволят значительно сократить сроки их проектирования и стоимость реализации.

Благодарности

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, соглашение № FSRF-2023-0003, "Фундаментальные основы построения помехозащищенных систем космической и спутниковой связи, относительной навигации, технического зрения и аэрокосмического мониторинга".

Литература

1. *Wang K., Li H., Zhang Q.* Parallel redundancy protocol for railway wireless data communication network // Wireless Communications and Mobile Computing, vol. 2022. 2022. С. 1-13. doi:10.1155/2022/3312569
2. *Zhang B., Yan B., Liu M., Xu A., Wang K., Wang Z.* Design of a highly reliable redundant network card supporting the IEEE1588 protocol // 2022 IEEE 5th International Conference on Electronics Technology (ICET). Chengdu, 2022. С. 830-836. doi:10.1109/ICET55676.2022.9824283.
3. *Bouwmeester J., Menicucci A., Gill E.K.A.* Improving CubeSat reliability: Subsystem redundancy or improved testing? // Reliability Engineering & System Safety, 2022. No. 220. С. 1-18. doi:10.1016/j.ress.2021.108288
4. *Xu J.* Topological structure and analysis of interconnection networks. Kluwer Academic Publishers, 2001. 350 с.
5. *Shooman M.L.* Reliability of computer systems and networks. Fault tolerance, analysis, and design. New York: Wiley , 2002. 551 с.
6. *Blokdyk G.* Network Redundancy. A Complete Guide Paperback, 2022. 307 с.
7. *Olenev V.L.* Analysis of requirements for modern spacecraft onboard network protocols. Информационно-управляющие системы, 2021. № 1. С. 8-16. doi:10.31799/1684-8853-2021-1-8-16
8. *Гаврилов М.А., Остиану В.М., Потехин А.И.* Надежность дискретных систем // Итоги науки. Серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика, 1969», ВИНИТИ, 1970. С. 7-104.
9. *Суворова Е.А., Шейнин Ю.Е.* Проектирование бортовых сетей SpaceFibre с пространственным резервированием // XLIV Академические чтения по космонавтике, посвященные памяти академика С.П. Королёва и других выдающихся отечественных ученых - пионеров освоения космического пространства : сборник тезисов (Москва, 28-31 января 2020 г.). Том 2. Москва, 2020. С. 226-228.
10. *Глухих М.И., Моисеев М.Ю., Егоров И.В., Крикун Т.С.* Автоматизация анализа надежности невосстанавливаемых информационно-управляющих систем // Информатика, телекоммуникации и управление, 2012. № 2. С. 81-90.
11. *Maurer A., Balagurin O., Greiner T., Herbst T., Kaiser T., Kayal H., Riegler C., Schwarz T.* Hardware and Software Redundancy Concepts onboard of SONATE-2 // Proceedings of 73rd International Astronautical Congress (IAC 2022), Paris, 2022.
12. *Шейнин Ю.Е., Оленев В.Л., Лавровская И.Я., Дымов Д.В., Кочура С.Г.* Разработка, анализ и проектирование транспортного протокола СТП-ИСС для бортовых космических сетей SpaceWire // Исследования наукограда. Красноярск, 2016. № 1-2. С. 21-30.
13. *Angelini P., Hanxleden R.* Graph drawing and network visualization // 30th International Symposium GD 2022. Tokyo, 2022. 499 с.
14. Cornejo A., Lynch N. Fault-tolerance through k-connectivity // Workshop on Network Science and Systems Issues in Multi-Robot Autonomy: ICRA, 2010. C.1-4.
15. *Сэджвик Н.* Фундаментальные алгоритмы на C++. Часть 5. Алгоритмы на графах: пер. с англ. СПБ: ООО «ДиаСофтЮП», 2002. 496 с.
16. *Saranya J.* A Study on point-to-point protocol in data communication and networking // International Journal of Computer Sciences and Engineering. 2019. № 7. С. 574-576. doi:10.26438/ijcse/v7i1.574576
17. *Harari F.* Graph Theory. Addison-Wesley Publishing Company, 1969. 274 с.
18. *Оленев В.Л., Шейнин Ю.Е., Лавровская И.Я., Коробков И.Л., Курбанов Л.И., Чумакова Н.Ю., Синёв Н.И.* Embedded Networks Design and Simulation // Tools and Technologies for the Development of Cyber-Physical Systems. Tampere: IGI Global, 2020. С. 77-118
19. *Cormen T.H., Leiserson C.E., Rivest R.L., Stein C.* Introduction to Algorithms (4th ed.). MIT Press and McGraw-Hill.2022. 1312 с.
20. Mastan S., Balakrishna U., Raju G., Kumar T.J. An articulation point based heuristic algorithm for minimum weight vertex cover problem // GIS science journal, Vol. 8, Issue 1, 2021. С. 1-9.
21. *Olenev V.L., Korobkov I.L., Chumakova N.Y., Sinyov N.I.* SANDS tool for design and simulation of onboard networks // 2021 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF). Saint-Petersburg, 2021. С. 1-8.

AUTOMATIC BUILDING OF FAULT-TOLERANT ON-BOARD NETWORKS

VALENTIN L. OLENEV
St. Petersburg, Russia, Valentin.Olenev@guap.ru

KEYWORDS: *fault-tolerance, structural redundancy, on-board networks, automation, network building*

ABSTRACT

Introduction: Fault tolerance is one of the key issues for on-board networks of autonomously operating vehicles. The number of devices in such networks is increasing; there is a need for methods for automating the design of networks, taking into account the required fault tolerance and technical limitations. **Purpose:** the purpose of the study is the implementation of methods and algorithms that contribute to the assessment of the fault tolerance of on-board networks, as well as the automatic generation of network structures with the required fault tolerance. **Methods:** The obtained methods and algorithms are based on the elements of graph theory, in particular – estimation of connectivity and shortest path search. The graphs are compiled based on the structure of the onboard network using a number of proposed rules. Structural redundancy is considered as a method of ensuring fault tolerance. The developed method provides reasonable addition of the network elements, in contrast to the often used copying of the network structure. The method consists of two stages: the new network structure is generated according to the required fault tolerance, and then iteratively refines the network structure in order to reduce hardware costs. Results: the use of the developed method simplifies and speeds up the design of on-board networks, since a formalized assessment of such a network is carried out at the design stage, before its actual assembly. The process of achieving of the required fault tolerance for large network structures takes less than a minute. **Practical relevance:** the presented method is implemented within the framework of the computer-aided system for design and modeling of onboard networks, a long-term approbation was carried out as part of real projects for the development of onboard networks of spacecraft. **Discussion:** there are no analogues of such a software implementation for on-board networks. The solutions obtained using the method in practice proved to be close to optimal; the research results have positive feedback from specialists.

REFERENCES

1. K. Wang, H.Li, Q. Zhang,"Parallel redundancy protocol for railway wireless data communication network," *Wireless Communications and Mobile Computing*, vol. 2022. 2022, pp. 1-13. doi:10.1155/2022/3312569

2. B. Zhang, B. Yan, M. Liu, A. Xu, K. Wang, Z. Wang, "Design of a highly reliable redundant network card supporting the IEEE1588 protocol. *2022 IEEE 5th International Conference on Electronics Technology (ICET)*," Chengdu, 2022, pp. 830-836. doi:10.1109/ICET55676.2022.9824283.

3. J. Bouwmeester, A. Menicucci, E. K. A. Gill, "Improving CubeSat reliability: Subsystem redundancy or improved testing?" *Reliability Engineering & System Safety*, 2022. No. 220, pp. 1-18. doi:10.1016/j.ress.2021.108288

4. J. Xu, "Topological structure and analysis of interconnection networks," Kluwer Academic Publishers, 2001. 350 p.

5. M.L. Shooman, "Reliability of computer systems and networks. Fault tolerance, analysis, and design," New York: Wiley , 2002. 551 p.

6. G. Blokdyk, "Network Redundancy," A Complete Guide Paperback, 2022. 307 p.

7. V.L. Olenev, "Analysis of requirements for modern spacecraft onboard network protocols," *Information and management systems*, 2021. No. 1, pp. 8-16. doi:10.31799/1684-8853-2021-1-8-16

8. M.A. Gavrilov, V.M. Ostianu , A.I. Potekhin, "Fault-tolerance of discrete systems," *The results of science. Series "Probability Theory. Math statistics. Theoretical cybernetics. 1969"*, VINITI, 1970, pp. 7-104.

9. E.A. Suvorova, Yu.E. Sheynin, "SpaceFibre Onboard Networks Design with Spatial Redundancy," *XLIV Academic readings on astronautics dedicated to the memory of Academician S.P. Korolev and other outstanding domestic scientists – pioneers of space exploration: a collection of abstracts*. Moscow, 28-31 January 2020. Vol. 2. Moscow, 2020, pp. 226-228.

10. M.I. Gluhih, M.Yu. Moiseev, I.V., Egorov, T.S. Krikun, "Automation of reliability analysis of non-restorable information and control systems," *Informatics, telecommunications and management*, 2012. No. 2, pp. 81-90.

11. A. Maurer, O. Balagurin, T. Greiner, T. Herbst, T. Kaiser, H. Kayal, C. Riegler, T. Schwarz, "Hardware and Software Redundancy Concepts on-board of SONATE-2," *Proceedings of 73rd International Astronautical Congress (IAC 2022)*, Paris, 2022.

12. Y.E. Sheynin, V.L. Olenev, I.Y. Lavrovskaya, D.V. Dymov, S.G. Kochura, "Development, analysis and design of the STP-ISS transport protocol for SpaceWire onboard networks," Krasnoyarsk, 2016. No. 1-2, pp. 21-30.

13. P. Angelini, R. Hanxleden, "Graph drawing and network visualization," *30th International Symposium GD* 2022. Tokyo, 2022. 499 p.

14. A. Cornejo, N. Lynch, "Fault-tolerance through k-connectivity," *Workshop on Network Science and Systems Issues in Multi-Robot* Autonomy: ICRA, 2010, pp.1-4.

15. R. Sedgewick, "Algorithms in C++ Part 5: Graph Algorithms," Addison-Wesley, 2002. 496 p.

16. J. Saranya, "A Study on point-to-point protocol in data communication and networking," *International Journal of Computer Sciences and Engineering*. 2019. No. 7, pp. 574-576. doi:10.26438/ijcse/v7i1.574576

17. Harari F. Graph Theory. Addison-Wesley Publishing Company, 1969. 274 p.

18. V.L. Olenev, Y.E. Sheynin, I.Y. Lavrovskaya, I.L. Korobkov, L.I. Kurbanov, N.Y. Chumakova, N.I. Sinyov, "Embedded Networks Design and Simulation," *Tools and Technologies for the Development of Cyber-Physical Systems*. Tampere: IGI Global, 2020, pp. 77-118.

19. T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, "Introduction to Algorithms," (4th ed.). MIT Press and McGraw-Hill, 2022. 1312 p.

20. S. Mastan, U. Balakrishna, G. Raju, T. J. Kumar, "An articulation point based heuristic algorithm for minimum weight vertex cover problem," *GIS science journal*. Vol. 8, Issue 1, 2021, pp. 1-9.

21. V.L. Olenev, I.L. Korobkov, N.Y Chumakova, N.I. Sinyov, "SANDS tool for design and simulation of onboard networks," *2021 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*. Saint-Petersburg, 2021, pp. 1-8.

INFORMATION ABOUT AUTHOR:

V.L. Olenev, PhD, docent, Chair of the Department of aerospace computer and software systems, Director of Aerospace R&D Centre of Saint Petersburg State University of Aerospace Instrumentation

For citation: V.L. Olenev. Automatic building of fault-tolerant on-board networks. *H&ES Reserch*. 2023. Vol. 15. No. 4. P. 4-13. doi: 10.36724/2409-5419-2023-15-4-4-13 (In Rus)

doi: 10.36724/2409-5419-2022-15-4-14-25

ОБЗОР КЛЮЧЕВЫХ ОСОБЕННОСТЕЙ ПРИ ПОСТРОЕНИИ ПОДВОДНОЙ ОПТИЧЕСКОЙ БЕСПРОВОДНОЙ СВЯЗИ

ПАВЛОВ
Иван Иванович [1]

ПАВЛОВА
Мария Сергеевна [2]

АБРАМОВА
Евгения Сергеевна [3]

АБРАМОВ
Сергей Степанович [4]

ЩЕРБАКОВ
Юрий Сергеевич [5]

Сведения об авторах:

[1] к.т.н., доцент, доцент кафедры "Радиотехнических устройств и техносферной безопасности", федеральное государственное бюджетное образовательное учреждение высшего образования "Сибирский государственный университет телекоммуникаций и информатики", г. Новосибирск, Россия, iipavlov02@mail.ru

[2] к.т.н., доцент кафедры "Радиотехнических устройств и техносферной безопасности", федеральное государственное бюджетное образовательное учреждение высшего образования "Сибирский государственный университет телекоммуникаций и информатики", г. Новосибирск, Россия, mspavlova@ngs.ru

[3] к.т.н., доцент, доцент кафедры "Радиотехнических устройств и техносферной безопасности", федеральное государственное бюджетное образовательное учреждение высшего образования "Сибирский государственный университет телекоммуникаций и информатики", г. Новосибирск, Россия, evgenka_252@mail.ru

[4] д.т.н., доцент, заведующий кафедрой "Радиотехнических устройств и техносферной безопасности", федеральное государственное бюджетное образовательное учреждение высшего образования "Сибирский государственный университет телекоммуникаций и информатики", г. Новосибирск, Россия, abramov@sibguti.ru

[5] к.т.н., доцент, доцент кафедры "Радиотехнических устройств и техносферной безопасности", федеральное государственное бюджетное образовательное учреждение высшего образования "Сибирский государственный университет телекоммуникаций и информатики", г. Новосибирск, Россия, ampal55@mail.ru

АННОТАЦИЯ

Введение: повышенный интерес к исследованию подводной среды океанов требуется для изучения флоры и фауны, дна океана, поиска полезных ископаемых, мониторинга существующих нефтяных вышек и других объектов в океане, сбора необходимой информации. Это все невозможно осуществить без качественной и надежной связи в подводных океанических условиях. Известными способами организации подводной беспроводной связи считаются системы связи с использованием акустических, радиочастотных и оптических волн. Зарубежные и российские исследователи отмечают, что подводная акустическая и радиочастотная беспроводная связь характеризуется небольшой скоростью передачи информации и высокой задержкой, но при этом позволяет передать информацию на большие расстояния. Потребность в высокой скорости передачи информации способствовала развитию подводной оптической беспроводной связи. Она позволяет организовать связь на большой скорости передачи информации с низкой задержкой на короткие расстояния. Увеличение дальности связи можно обеспечить при помощи сетевых технологий построения таких систем связи. **Цель исследования:** целью исследования является обзор научных работ зарубежных исследователей в анализе теоретических и экспериментальных исследований в области подводной оптической беспроводной связи. **Результаты:** Показано, что подводная оптическая беспроводная связь может быть организована непосредственно между надводной станцией и оптическими узлами или же с использованием сетевых технологий построения сети может организовать многоуровневую систему связи от надводной станции до оптических узлов с применением оптических точек доступа и/или оптических базовых станций. Проведенный анализ показал, что при организации подводной оптической беспроводной связи основным влиянием на качество и надежность связи непосредственно оказывает подводная океаническая среда. **Практическая значимость работы** заключается в систематизации научных работ зарубежных исследователей по проблемам подводной оптической беспроводной связи и использовании в качестве основы для будущих прикладных исследований.

КЛЮЧЕВЫЕ СЛОВА: *оптический беспроводной канал связи, подводная оптическая беспроводная связь, скорость передачи информации, дальность связи, световой пучок лазера, тип океанической воды, модуляция интенсивности; когерентная модуляция*

Для цитирования: *Павлов И.И., Павлова М.С., Абрамова Е.С., Абрамов С.С., Щербаков Ю.С.* Обзор ключевых особенностей при построении подводной оптической беспроводной связи // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 14-25. doi: 10.36724/2409-5419-2023-15-4-14-25

Введение

Изучением планеты Земля во всем мире занимаются миллионы людей. Еще много лет назад людей интересовало почему светит солнце, идет дождь, некоторые животные могут жить на суше, но при этом не могут плавать в воде и многое другое. Это способствовало в дальнейшем развитию различных видов научных исследований.

В настоящее время большое внимание уделяется изучению океанов. Если учесть тот факт, что поверхность Земли покрыта водой на 70,8 % и при этом общий размер водной поверхности составляет примерно 510 000 000 км^2, то из этого количества воды где-то 360 000 000 км^2 составляют воды океанов. Поэтому высокий интерес к исследованию подводной среды океанов необходим для изучения животного и растительного мира, исследования дна океана, поиска полезных ископаемых в недрах океана, мониторинга существующих нефтяных вышек и других объектов в океане, наблюдение и сбор различной информации для изучения океанов. Для осуществления всех перечисленных операций необходима качественная и надежная связь из подводной среды океанов с поверхностными станциями.

Наиболее распространенными видами подводной беспроводной связи (рис. 1) являются системы связи, лежащие на основе акустических волн, радиочастотных волн и оптических волн.

Одной из первых и распространенных технологий подводной беспроводной связи была подводная акустическая беспроводная связь, которая позволяет передавать информацию от передатчика к приемнику на очень большие расстояния. В работе [1] еще в 1995 году была представлена система подводной акустической беспроводной связи со скоростью передачи информации 40 кбит/с.

Уже через год в 1996 году ученные разработали систему подводной акустической беспроводной связи, которая работала на скорости 8 кбит/с, но при этом позволяла передавать информацию в подводной среде океанов на расстояние 13 км в горизонте и на 20 м в глубину (по вертикали) от передатчика сообщений к получателю [2].

Рис. 1. Основные виды подводной беспроводной связи

Развитие подводной акустической беспроводной связи на этом не остановилось и уже в 2005 году в работе [3] была разработана система, в которой увеличили скорость передачи информации до 125 кбит/с и использовался метод 32-QAM с коэффициентом символьных ошибок 10^{-4}. Система подводной акустической беспроводной связи имеет много недостатков, например таких, как высокая задержка при передаче информации из-за низкой скорости распространения, большое рассеяние, низкая пропускная способность и негативное влияние акустических волн на подводных млекопитающих, рыб и растений.

Ключевым недостатком подводной акустической беспроводной связи является очень низкая скорость передачи информации, что привело к развитию и исследованию подводных беспроводных систем на основе низкочастотных радиоволн. Если рассмотреть работы зарубежных авторов, то в [4] была предложена подводная беспроводная связь на основе микроволн над поверхностью океана и в результате передачу информации можно было осуществить на несколько десятков километров. Этот же подход на микроволнах был использован в работе [5], где скорость передачи информации в подводной среде океана была увеличена до 500 кбит/с и передавалась на расстояние до 90 м по горизонтали. В работе [6] авторам удалось увеличить пропускную способность подводной беспроводной связи на основе микроволн до 10 Мбит/с и с возможностью передачи по горизонтали на расстояние 100 м.

Тем не менее подводная радиочастотная беспроводная связь имеет большой недостаток в виде серьезного затухания в воде. Также для осуществления сеанса связи потребуются большие антенны и ограничение прибрежными участками океана.

Однако организовать работу подводной радиочастотной беспроводной связи возможно, если использовать сверхнизкие частоты. Это позволит уменьшить уровень затухания, но «платой» за это будет высокая стоимость оборудования и низкая скорость передачи информации.

Низкая скорость передачи информации и низкая пропускная способность подводных акустических и радиочастотных беспроводных систем послужила развитию подводной беспроводной связи на основе оптических волн.

Многочисленные исследования и эксперименты доказали, что затухания оптических волн в подводной среде океана в диапазоне длин волн 450-550 нм – это синий и зеленый свет, являются наименьшими относительно других длин волн. Все исследования в изучении подводной оптической беспроводной связи заключались в увеличении скорости передачи информации [7].

На сегодняшний день наиболее распространенной подводной оптической беспроводной связью является связь на основе лазера. В работе [8] была рассмотрена подводная оптическая беспроводная связь на основе зеленого лазера с длиной волны 532 нм, где в результате получили скорость передачи информации равной 1 Гбит/с, но расстояние составляло всего 2 м. В работе [9], вышедшей в 2015 году, авторами был использован синий лазер с длиной волны 405 нм, чем добились скорости передачи информации до 1,45 Гбит/с и передачи на расстояние 4,8 м.

Подводная оптическая беспроводная связь имеет преимущество в высокой скорости передачи информации, низкой задержки и энергоэффективности из-за небольшого расстояния передачи информации.

Для увеличения расстояния передачи информации от передатчика сообщений к получателю в подводной оптической беспроводной связи необходимо использовать сетевые решения при построении таких систем.

В данной статье рассмотрим возможную архитектуру для подводной оптической беспроводной связи, а также основные свойства оптических волн при распространении в подводной среде океана и методы модуляции в подводной оптической беспроводной связи.

Архитектура подводной оптической беспроводной связи

Подводная оптическая беспроводная связь может работать в двух режимах. Первый режим, режим прямого подключения (ad-hoc) – это самый распространенный тип беспроводной сети, в котором нет зависимости от установленного сетевого оборудования.

В этом режиме не предусмотрен центральный блок управления, вследствии чего передача информации осуществляется непосредственно оптическими узлами. Путь прохождения информации от источника сообщения до получателя формируется в момент подключения и является динамическим. Поэтому в данной сети для оптических узлов необходимы навыки самоорганизации и самонастройки.

Второй режим – это режим инфраструктуры. Он строится на оптических точках доступа, которые распространяют оптические волны в разные стороны, или же возможно использование оптических базовых станций. Оптические точки доступа и/или оптические базовые станции организуют подводную локальную сеть, в которой каждая оптическая точка доступа или оптическая базовая станция обслуживает и координирует работу в зоне своего покрытия.

На рисунке 2 представлен пример организации подводной оптической связи. Для организации связи между подводными оптическими узлами используют оптические каналы темно-зеленого цвета. Связь от оптической базовой станции до подводного оптического узла осуществляется оптическими каналами оранжевого цвета.

Если необходимо связаться оптическим базовым станциям на одной глубине, то используются горизонтальные оптические каналы красного цвета, а связь между оптическими базовыми станциями на разных глубинах осуществляется вертикальными оптическими каналами синего цвета. Для увеличения энергоэффективности данной системы возможно организовать питание оптических узлов от солнечных батарей, которые расположены на поверхностных буях и подключенных к оптическим узлам по проводному каналу.

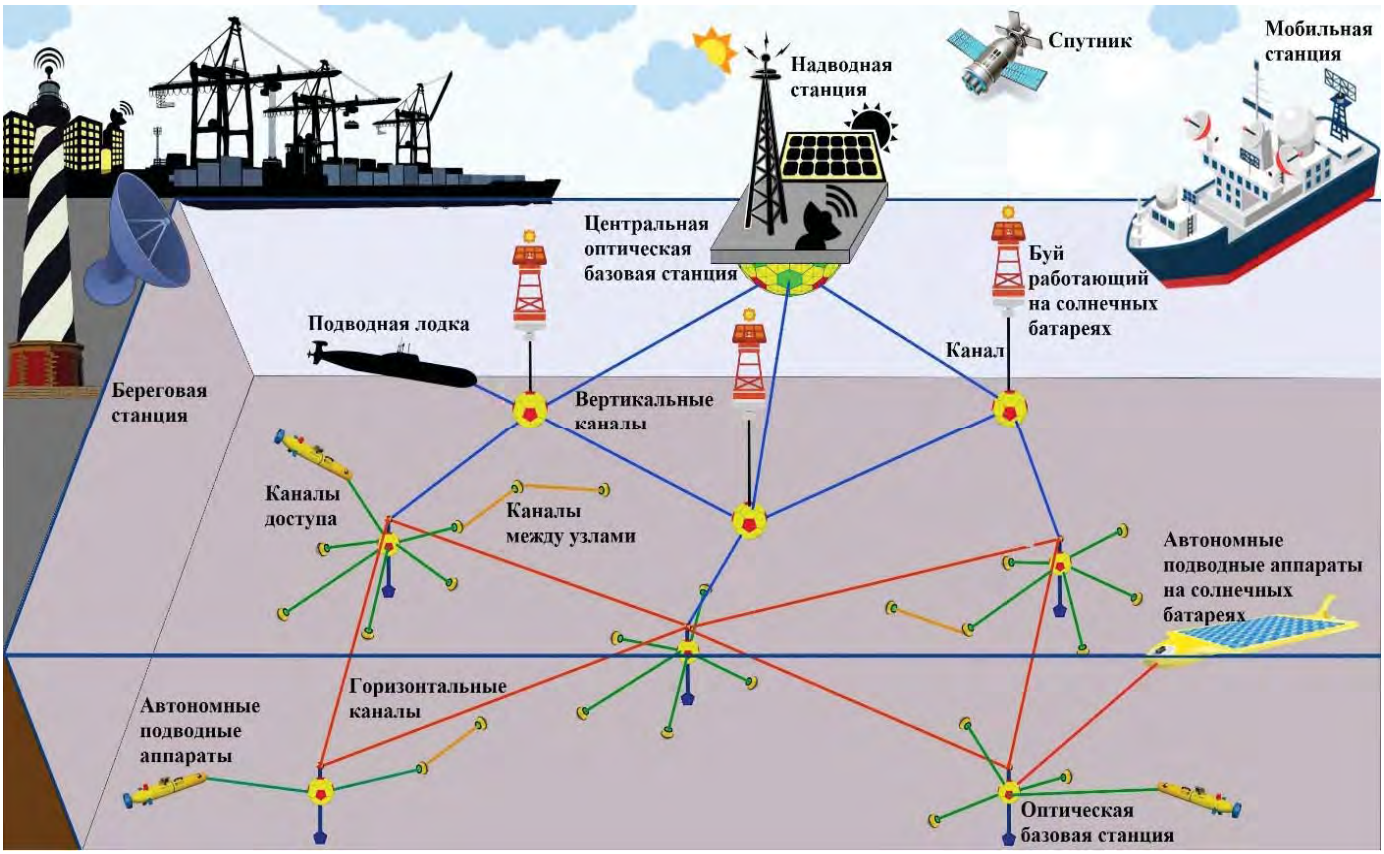


Рис. 2. Пример организации подводной оптической беспроводной связи

Все подводные лодки и автономные подводные аппараты могут организовать сеансы связи между собой и с наземными станциями, находящимися в одной подводной оптической беспроводной связи. Вся необходимая информация, собранная подводными оптическими узлами, поступает через центральную оптическую базовую станцию на надводную станцию. Далее эта информация может быть передана с помощью радиочастотных сетей или спутниковой связи на мобильную станцию или береговую станцию.

Рис. 3. Архитектура подводной оптической беспроводной связи

На рисунке 3 представлена одна из возможных архитектур подводной оптической беспроводной связи. Более подробно рассмотрим принцип пространственного покрытия подводной оптической беспроводной связи.

1. Стационарная одномерная подводная оптическая беспроводная связь организуется за счет того, что оптические узлы образуют горизонтальную линию, которая может находиться у поверхности воды, так как каждый оптический узел закреплен за поверхностным буем или же опущен на океаническое дно. При этом оптические узлы находятся в стационарном состоянии, т.е. не могут передвигаться в пространстве. В этом случае вся информация передается непосредственно от оптического узла на надводную станцию. Такая архитектура подводной оптической беспроводной связи напоминает топологию «звезды».

2. Подвижная одномерная подводная оптическая беспроводная связь организована, как и стационарная одномерная подводная оптическая беспроводная связь, где оптические узлы работают автономно. При этом оптический узел, закрепленный за поверхностным буем, может перемещаться по горизонтали и собирать необходимую информацию, а затем передавать ее на надводную станцию. Оптические узлы, которые размещаются на океаническом дне, также могут передвигаться по горизонтали и после сбора информации поднимаются к поверхности для передачи собранной информации на надводную станцию.

3. Стационарная двумерная подводная оптическая беспроводная связь представляет собой группу оптических узлов, которые подключены к оптической базовой станции. Вся информация, собранная оптическими узлами, передается на оптическую базовую станцию и уже с оптической базовой станции поступает на надводную станцию. В стационарной двумерной подводной оптической беспроводной связи все оптические узлы связываются с оптической базовой станцией по горизонтальным оптическим каналам. А оптическая базовая станция связывается с надводной станцией по вертикальным оптическим каналам. В рассматриваемой стационарной двумерной подводной оптической беспроводной связи предусматривается то, что все оптические узлы находятся на одной глубине в одной горизонтальной плоскости. Такое построение сети возможно звездой, кольцом, сотой или ячеистой топологией.

4. Подвижная двумерная подводная оптическая беспроводная связь строится по аналогии со стационарной двумерной подводной оптической беспроводной связью, но в данной модели оптические узлы могут свободно плавать в подводной среде океана. Кроме того, оптическая базовая станция тоже может передвигаться и собирать информацию с различных подводных оптических узлов или датчиков. Подвижная двумерная подводная оптическая беспроводная связь является более сложной динамической топологией по сравнению со стационарной двумерной подводной оптической беспроводной связью.

5. Стационарная трехмерная подводная оптическая беспроводная связь организована за счет размещения оптических узлов на разной глубине. Поэтому для связи между надводной станцией и оптическим узлом необходимо организовать три уровня связи.

а) Связь между оптическими узлами на разной глубине или связь между оптическими узлами и оптической подводной станцией на одной глубине.

б) Связь между оптическими узлами и оптической базовой станцией на разных глубинах или связь между оптическими базовыми станциями на одной глубине.

в) Связь оптической базовой станции с надводной станцией.

Все оптические узлы и оптические базовые станции не могут передвигаться как по горизонтали, так и по вертикали.

6. Подвижная трехмерная подводная оптическая беспроводная связь строится на основе автономных подводных аппаратов или дистанционно управляемых подводных аппаратов, которые могут передвигаться на разную глубину и в разных направлениях. Это позволяет увеличить производительность подводной оптической беспроводной связи.

Основные положения подводной оптической беспроводной связи на физическом уровне

Физический уровень подводной оптической беспроводной связи, как любой проводной сети, включает в себя различные основные функции связи, такие как обработка сигнала, модуляция и демодуляция сигнала, кодирование сигнала, моделирование и оценка канала связи и многое другое.

Как уже упоминалось выше подводная оптическая беспроводная связь по сравнению с подводной беспроводной связью на основе акустических и радиочастотных волн позволяет организовать канал связи с высокой скоростью передачи информации на расстояния в десятки метров с очень низкой задержкой. Также для создания приемопередающих устройств подводной оптической беспроводной связи могут быть использованы недорогие, энергосберегающие лазеры и

при этом размер самого устройства может быть небольшим. Так как связь между оптическими узлами или между оптическим узлом и оптической базовой станцией осуществляется в режиме «точка-точка», то она еще и обеспечивает повышенную надежность передачи информации.

Но несмотря на все это у подводной оптической беспроводной связи существует множество проблем, которые необходимо устранять. Например, при организации оптической беспроводной связи между двумя узлами из-за смещения приемопередатчиков в пространстве может произойти разъединение. Данное разъединение связано с движением глубинных течений подводной среды океана в зависимости от глубины и поверхностных движений океанической воды. Также если в качестве несущей волны будет выбран лазер синего или зеленого цвета, он все равно будет подвержен поглощению, рассеиванию из-за воздействия молекул воды и твердых частиц, находящихся в океанической воде, на фотоны. Этот факт приводит к ухудшению производительности и снижению дальности связи.

Поэтому для увеличения дальности связи используется сетевая схема построения подводной оптической беспроводной связи. Для устранения несоосности оптических узлов требуются очень точные алгоритмы выравнивания, которые должны поддерживать связь между оптическими узлами постоянно.

Поэтому при исследовании подводной оптической беспроводной связи основной аспект ложится на изучение физического уровня. Сюда можно отнести распространение оптических волн в подводной среде океана, ретрансляцию, маршрутизацию, развертывание, подвижность оптических узлов, самонастройку, самоорганизацию и т.д.

Основные свойства оптических волн при распространении в подводной среде океана

Распространение оптических волн в подводной среде океана разнообразно и очень сильно зависит от того, где организован оптический беспроводной канал связи, на какой глубине и какую физико-химическую структуру имеет океан в данном месте. Для того, чтобы организовать оптический беспроводной канал связи на физическом уровне необходимо рассмотреть оптические свойства подводной среды океана, которые делятся на присущие и видимые.

К присущим оптическим свойствам можно отнести коэффициент поглощения, коэффициент рассеяния и коэффициент затухания. Все эти коэффициенты являются сильно зависящими от химического состава океанической воды.

Видимые оптические свойства океанической подводной среды основаны на таких факторах, как среда и геометрическая структура световых пучков, которая включает в себя коэффициент отражения, коэффициент излучения и проводимость светового пучка лазера.

При передаче информации по оптическому беспроводному каналу связи большое влияние на дальность оказывает поглощение подводной среды океана. Это связано с тем, что мощность сигнала излучаемого светового пучка при распространении постоянно уменьшается из-за физических свойств подводной среды океана. Кроме поглощения на мощность сигнала оказывает такое свойство подводной среды океана, как рассеяние светового пучка. Фотоны светового пучка рассеиваются в разных направлениях и в конечном результате приемником принимается не весь переданный оптический сигнал из-за ограниченных размеров апертуры приемника.

Оставшиеся фотоны светового пучка могут быть приняты с задержкой из-за более длинных и различных путей передачи части светового пучка лазера вследствие рассеяния. Этот фактор приводит к тому, что возникает межсимвольная интерференция, многопутевое замирание и временное дрожание.

В работе [10] авторами предложена геометрическая модель прохождения светового пучка через водную среду океана (рис. 4).

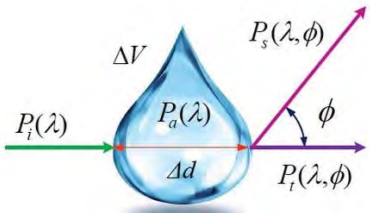


Рис. 4. Геометрическая модель прохождения светового пучка через водную среду океана

На основе данной геометрической модели прохождения светового пучка через водную среду океана можно сформулировать такие понятия, как коэффициент поглощения и коэффициент рассеяния.

Рассматриваемая модель имеет какой-то объем воды ΔV с толщиной Δd, через который проходит световой пучок лазера с длиной волны λ и полной мощностью сигнала $P_i(\lambda)$. Проходя через объем воды часть мощности сигнала $P_i(\lambda)$ поглощается телом воды:

$$P_a(\lambda) = \alpha(\lambda)\cdot P_i(\lambda),$$

а часть мощности сигнала $P_i(\lambda)$ рассеивается из-за изменения направления:

$$P_s(\lambda, \Phi) = \beta(\lambda)\cdot P_i(\lambda),$$

где $\alpha(\lambda)$ и $\beta(\lambda)$ – коэффициент поглощения и коэффициент рассеяния, соответственно,

Φ – телесный угол рассеяния светового пучка лазера.

Оставшаяся мощность сигнала продолжает распространяться по заданной траектории:

$$P_t(\lambda, \Phi) = \gamma(\lambda)\cdot P_i(\lambda),$$

где $\gamma(\lambda)$ – коэффициент остаточной мощности сигнала.

Согласно закону сохранения мощности коэффициент поглощения, коэффициент рассеяния и коэффициент остаточной мощности сигнала связаны между собой следующим соотношением:

$$\alpha(\lambda) + \beta(\lambda) + \gamma(\lambda) = 1.$$

Исходя из рассматриваемой модели определили спектральную поглощающую способность $A(\lambda)$ как отношение поглощаемой мощности сигнала к полной мощности сигнала:

$$A(\lambda) = \frac{P_a(\lambda)}{P_i(\lambda)}.$$

Спектральное рассеяние *B(λ)* определяется как отношение рассеянной мощности сигнала к полной мощности сигнала:

$$B(\lambda) = \frac{P_s(\lambda)}{P_i(\lambda)}.$$

Коэффициент поглощения *α(λ)* определяется путем взятия предела, когда толщина *Δd* приближается к нулю, для отношения спектрального поглощения *A(λ)* к толщине *Δd* следующим образом:

$$\alpha(\lambda) = \lim_{\Delta d \to 0} \frac{A(\lambda)}{\Delta d} \quad \left(m^{-1}\right).$$

Коэффициент рассеяния *β(λ)* определяется путем взятия предела, когда толщина *Δd* приближается к нулю, для отношения спектрального рассеяния *B(λ)* к толщине *Δd* следующим образом:

$$\beta(\lambda) = \lim_{\Delta d \to 0} \frac{B(\lambda)}{\Delta d} \quad \left(m^{-1}\right).$$

Сумма коэффициента поглощения *α(λ)* и коэффициента рассеяния *β(λ)* в результате дадут коэффициент затухания мощности сигнала в подводной среде океана:

$$c(\lambda) = \alpha(\lambda) + \beta(\lambda).$$

Коэффициент экстинкции *c(λ)* показывает ослабление мощности сигнала и сильно зависит на от глубины нахождения оптического беспроводного канала связи и от типа океанической воды, в которой распространяется световой пучок лазера.

Встречаются три наиболее распространенных типа океанической воды.

1. Чистая океаническая вода состоит из молекул чистой воды (H_2O) и растворенных в ней солей, таких как *NaCe*, $MgCe_2$, Na_2SO_4, *HCe* и многих других. Суммарное поглощение молекул чистой воды и растворенных в ней солей составляет общее поглощение чистой океанической воды и является небольшой величиной. Также стоит отметить, что коэффициент рассеяния в чистой океанической воде тоже имеет небольшие значения и поэтому световой пучок лазера распространяется по заданной траектории с малой величиной дисперсий.

2. Прибрежная океаническая вода включает в себя большую концентрацию различных растворенных частиц. Из-за этого в прибрежной океанической воде коэффициент поглощения и коэффициент рассеяния имеют более высокие значения, что негативно влияет на передачу светового пучка лазера в созданном оптическом беспроводном канале связи.

3. Мутная вода гавани является наиболее неблагоприятной средой для распространения оптических волн, так как в ней самая наибольшая концентрация взвешенных и цветных растворенных органических частиц, которые увеличивают значения коэффициента поглощения и коэффициента рассеяния практически до максимальных.

На распространение оптических волн в подводной среде океана также особую роль играет глубина погружения оптического беспроводного канала связи. И в зависимости от того, на сколько глубоко солнечный свет проникает в океан, различают несколько вертикальных зон разделения подводной среды океана (рис. 5).

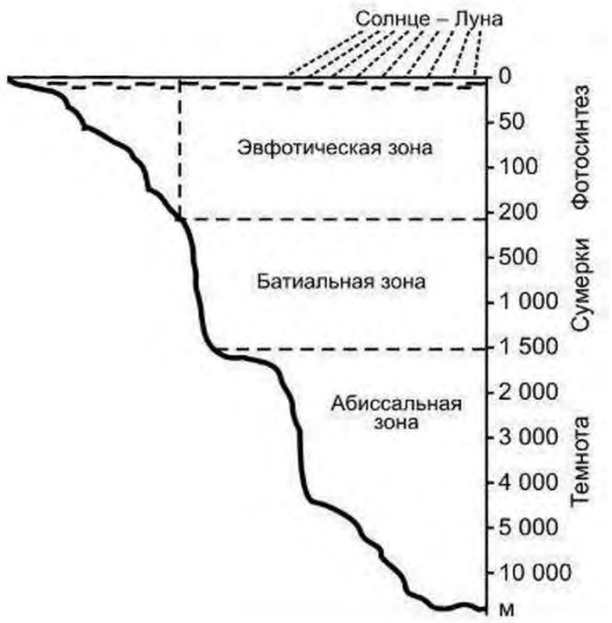


Рис. 5. Вертикальное зонирование подводной среды океана

Самый верхний слой океана, начинающийся от поверхности океана, называется фотической зоной, она располагается до той глубины, куда проникает солнечный свет. Фотическая зона состоит из двух зон. Первая зона называется эвфотической зоной и располагается в среднем до 200 м от поверхности океана вглубь. В этой зоне довольно достаточная освещенность для поддержания фотосинтеза и жизнедеятельности океанической флоры и фауны.

Вторая зона – это батиальная зона, она располагается в глубине океана от 200 м до 1500 м. Ее еще называют сумеречной зоной, в ней освещенность солнечным светом падает и не может быть достаточной для фотосинтеза.

Самой последней зоной является абиссальная зона, она располагается на глубине ниже 1500 м. Ее еще называют зоной вечной темноты, солнечный свет в эту зону не попадает. Воды в этой зоне отличаются небольшой подвижностью, низкой температурой и низким содержанием биогенных веществ.

Несмотря на то, что фотическая зона располагается в верхних слоях океана, а при этом средняя глубина океана составляет примерно от 3,8-4,3 км, фотическая зона представляет собой самую большую часть океана, поэтому коэффициент затухания в верхних слоях до 100 м начинается с величины 0,05 $м^{-1}$, далее на глубине около 100 м он достигает 0,1 $м^{-1}$ и, чем глубже находится оптический беспроводной канал связи, тем больше коэффициент затухания.

Океаническая турбулентность определяется как хаотические быстрые неупорядоченные изменения показателя преломления из-за колебаний океанических свойств подводной среды, таких как скорость, давление, соленость, плотность, температура воды [11]. Океаническая турбулентность приводит к непостоянству приема передаваемого светового пучка лазера и, соответственно, к значительному снижению производительности оптического беспроводного канала связи.

Наиболее важной задачей при организации оптической беспроводной связи между оптическими беспроводными приемопередатчиками является организация и поддержание постоянной надежной и качественной связи. Это достигается с помощью наведения и выравнивания, т.е. устранения несоосности оптических беспроводных приемопередатчиков.

В основном при наведении и выравнивании оптических беспроводных приемопередатчиков основными проблемами являются, во-первых, смещение центра светового пучка лазера передатчика и центра апертуры приемника на некоторый телесный угол. Этот фактор может быть вызван неточной информацией о месторасположении приемника. И во-вторых, возможны дрожания светового пучка лазера, которые могут возникнуть из-за случайного смещения центра светового пучка лазера передатчика и центра апертуры приемника в результате океанической турбулентности, глубинных течений и колебаний в зависимости от глубины, а также от случайных движений океанической флоры и фауны.

Если первую проблему наведения и выравнивания можно решить с помощью более точной информации о месторасположении приемника и наиболее эффективных алгоритмов расчета местоположения приемника, то вторая проблема остается трудно решаемой, так как контролировать случайные процессы океанической подводной среды очень тяжело.

Однако в работе [12] авторы рассмотрели то, что если оптический беспроводной канал связи находится в прибрежной океанической воде или мутной воде гавани жесткие требования, предъявляемые к наведению и выравниванию, становятся слабее из-за высокой дисперсии светового пучка лазера.

При прохождении светового пучка лазера через оптический беспроводной канал связи присутствует эффект рассеяния фотонов в результате свойств океанической подводной среды. В связи с чем часть светового пучка лазера может отклониться от заданной траектории и следовать по другому пути распространения и с большим расстоянием. Данная часть светового пучка лазера поступает на апертуру приемника в разные моменты времени, что приводит к задержке или временной дисперсии, а также к межсимвольной интерференции.

Многолучевая задержка светового пучка лазера в основном возникает на мелководье из-за отражения от поверхности океана, океанического дна и препятствия в непосредственной близости. А межсимвольная интерференция проявляет себя при большой скорости передачи информации по оптическим беспроводным каналам связи. Авторами в работе [13] был проведен анализ и получен результат, что при расстоянии 50 метров для поляризованного светового пучка лазера со скоростью передачи информации 1 Гбит/с влияние межсимвольной интерференции является существенной. В работе [14] было доказано обратное, что при использовании метода Монте-Карло влияние межсимвольной интерференции на коротких расстояниях не существенно.

Основные способы модуляции в подводной оптической беспроводной связи

В подводной оптической беспроводной связи можно выделить два основных типа модуляции: модуляция интенсивности и когерентная модуляция.

Модуляция интенсивности, или же некогерентная модуляция, представляет собой вид модуляции, где выходная оптическая мощность светового пучка лазера изменяется в соответствии с определенной характеристикой модулированного сигнала.

Организовать модуляцию в подводной оптической беспроводной связи можно двумя способами – использовать встроенный модулятор или внешний модулятор. Если используется встроенный модулятор, то передача «1» и «0» происходит за счет включения и выключения источника света, т.е. используется ток источника света. При передаче «1» источник света включается, а при передаче «0» источник света выключается.

Встроенный модулятор отличается невысокой ценой и простотой сборки, но при этом он ограничен в дальности и скорости передачи информации. Это связано с тем, что возникает эффект чирпинга (*chirping effect*) – наложение паразитной модуляции длины волны генерации, созданной импульсным преобразованием лазера.

Во внешних модуляторах источник света не выключается для непрерывной передачи светового пучка лазера. Чтобы передать сообщение внешний модулятор модулирует сигнал за счет изменения интенсивности и фазы сигнала, что в свою очередь обеспечивает блокировку и прохождение светового пучка лазера к приемнику. Внешние модуляторы отличаются очень высокой скоростью передачи информации и большой дальностью связи. Это обеспечивается за счет постоянной мощности передачи и высокой скорости переключения. Но это приводит к увеличению стоимости оборудования и сложности конструкции, а также не эффективному использованию мощности источника света.

При использовании в схеме встроенного модулятора и модуляции интенсивности такая система называется *IM/DD*-модуляцией. Она является самой распространенной, так как у нее низкая стоимость, простота изготовления и в ней не надо хранить информацию о фазе.

Одним из видов *IM/DD*-модуляции является схема *ON/OFF Keying* (*OOK*), что в переводе с английского означает управление включением/выключением. В схеме *OOK* «1» и «0» передаются через оптический беспроводной канал связи наличием или отсутствием светового пучка лазера. В схеме *OOK* организованы форматы импульсов с возвратом к нулю или без возврата к нулю. Если используется формат импульса без возврата к нулю, то для передачи «1» необходимо использовать всю длительность бита, а в формате импульса с возвратом к нулю используется только часть длительности бита.

Схема *OOK* очень сильно зависит от изменения оптического беспроводного канала связи и поэтому производительность данной схемы сильно ухудшается.

Для того чтобы улучшить производительность схемы *OOK*, необходимо использовать механизм динамического порога обнаружения ошибок. В работе [15] представлен данный механизм динамического обнаружения ошибок в соответствии с оценкой состояния оптического беспроводного канала связи.

Схема *OOK* стала очень популярной и практичной за счет низкого энергопотребления, эффективного использования полосы пропускания и простоты конструкции. В работах [16, 17] можно ознакомиться с теоретическими и экспериментальными исследованиями для подводной оптической беспроводной связи.

Позиционно-импульсная модуляция – также одна из самых широко используемых технологий в подводной оптической беспроводной связи. В данной модуляции каждый из *M* переданных битов в виде импульса в течение 2^M временных интервалов, положение которых соответствует отправленному сообщению.

Позиционно-импульсная модуляция позволяет обеспечить более высокую мощность и спектральную эффективность в обмен на более сложный приемопередатчик. Но из-за эффекта дрожания предъявляются жесткие требования к синхронизации, что приводит к снижению производительности. Поэтому традиционно-импульсная модуляция была усовершенствована и получены следующие виды модуляции: дифференциальная позиционно-импульсная модуляция [18], дифференциальная амплитудная позиционно-импульсная модуляция [19] и многие другие. Более подробно можно ознакомиться с аналитическими и экспериментальными исследованиями позиционно-импульсной модуляции для подводной оптической беспроводной связи в работах авторов [20 - 27].

Широтно-импульсная модуляция позволяет уменьшить среднюю мощность светового пучка лазера разделением его на дискретные части. Поэтому широтно-импульсная модуляция снижает пиковую мощность передачи путем распределения общей мощности на *M* временных интервалов, что обеспечивает более высокую среднюю мощность с увеличением *M* временных интервалов. Широтно-импульсная модуляция главным образом выгодна из-за своей спектральной эффективности и невосприимчивостью к межсимвольной интерференции.

В цифровой импульсно-интервальной модуляции за импульсами «Вкл» следуют временные интервалы «Выкл», количество которых равно десятичному значению символа передаваемой информации. Более подробно можно ознакомиться в работе [28].

Цифровая импульсно-интервальная модуляция является асинхронным методом модуляции, который также может поддерживать переменную длину символов. Несмотря на то, что цифровая импульсно-интервальная модуляция обеспечивает высокую мощность и спектральную эффективность, она подвержена ошибочному приему информации во время процедуры демодуляции. Более подробно о цифровой импульсно-интервальной модуляции для подводной оптической беспроводной связи можно ознакомиться в работах [29-31].

В отличие от модуляции интенсивности в когерентной модуляции для передачи информации используется как амплитудная, так и фазовая модуляция. В приемнике задающий генератор преобразует оптическую несущую в базовую полосу или радиочастотную промежуточную частоту, которая называется гомодинным и гетеродинным преобразованием, соответственно.

Когерентная модуляция обеспечивает более высокую чувствительность приемника, спектральную эффективность и устойчивость к фоновому шуму, но с дополнительными затратами и сложностью. Для более подробного рассмотрения когерентной модуляции в подводной оптической связи можно ознакомиться с работами [22, 32-34].

В конце статьи хотелось бы уделить внимание источникам шума в оптических беспроводных приемопередатчиках. В работе [35] выявлены следующие источники шума: это темновой ток фотодиода, шум передатчика, дробовой (пуассоновский) шум, тепловой шум и фоновый шум. Можно отметить то, что на практике темновой ток фотодиода пренебрежительно мал, его влияние можно не учитывать. На передаваемый световой пучок лазера влияет шум передатчика, вызванный колебаниями интенсивности света. Шум передатчика обычно рассчитывается относительным шумом лазера, который оказывает незначительное влияние на производительность приемника [36]. Тепловой шум обычно рассчитывается как гауссовский случайный процесс с нулевым средним значением, который является результатом поведения электронных схем, особенно нагрузочного резистора. Дробовой шум рассчитывается как пуассоновский процесс, который возникает из-за случайных флуктуаций тока фотодиода. Если количество принимаемых фотонов велико, пуассоновский процесс может быть аппроксимирован гауссовским процессом, как для приемников на основе *PIN*-диодов, так и для приемников на основе лавинных фотодиодов [36, 37].

Фоновый шум сильно зависит от типа воды, глубины расположения оптического беспроводного канала связи и длины волны оптической несущей. В эвфотической зоне солнечная интерференция может рассматриваться как основной источник фонового шума.

Заключение

Проведенный обзор и анализ научных работ зарубежных исследователей показал, что к основным способам организации подводной беспроводной связи можно отнести системы связи с применением акустических, радиочастотных и оптических волн.

Преимуществом подводных акустических и радиочастотных беспроводных сетей является передача информации на дальние расстояния, но при этом они обладают низкой скоростью передачи информации и высокой задержкой. Все эти недостатки компенсирует в себе подводная оптическая беспроводная связь. Она позволяет обеспечить высокие скорости передачи информации с низкой задержкой, а за счет построения по сетевым технологиям – увеличить дальность связи.

Предложенная архитектура подводной оптической беспроводной связи может работать в двух режимах. В первом режиме надводная станция непосредственно устанавливает связь с оптическими узлами. Во втором режиме строится, в

основном, многоуровневая система связи за счет использования оптических точек доступа и/или оптических базовых станций.

Рассмотренные научные исследования показали, что основное и большое влияние на качество и надежность связи оказывает подводная океаническая среда, где находится оптический беспроводной канал связи. Выделяют присущие оптические свойства подводной океанической среды, к которой относятся коэффициенты поглощения, рассеяния и затухания. Все они сильно зависят от химического состава воды океана. И видимые оптические свойства – это такие факторы, как среда и геометрическая структура светового пучка лазера. К ним относятся коэффициенты отражения, излучения и проводимости светового пучка лазера. Каждый из оптических свойств по-своему влияет на источник света и оказывает неблагоприятное воздействие на него.

Показана возможность использования модуляции интенсивности и когерентной модуляции в подводной оптической беспроводной связи. Это позволяет использовать два вида модуляторов, такие как встроенный и внешний модуляторы.

Встроенный модулятор простой в сборке и имеет низкую стоимость, но при этом не обеспечивает связь на дальние расстояния с высокой скоростью передачи информации. Внешний модулятор позволяет обеспечить высокую скорость передачи информации на большие расстояния, но это приводит к увеличению стоимости оборудования, сложности конструкции и неэффективности использования мощности источника света.

Подводя итог, можно сказать, что систематизация научных работ зарубежных исследователей по подводной оптической беспроводной связи выявила достаточно много проблемных вопросов, которые могут послужить основой для будущих прикладных исследований.

Литература

1. *Стоянович М.* Последние достижения в области высокоскоростной подводной акустической связи // IEEE Журнал океанической инженерии, том 21, № 2. С. 125-136, Апрель 1996, DOI:10.1109/48.486787

2. *Зелински А., Юн Я.Х., Ву Л.* Анализ эффективности цифровой акустической связи в мелководном канале // IEEE Журнал океанической инженерии, том 20, № 4. С. 293-299, Октябрь 1995, DOI:10.1109/48.468244.

3. *Очи Х., Ватанабе Е., Шимура Т.* Базовое исследование подводной акустической связи с использованием 32-квадратурной амплитудной модуляции // Японский журнал прикладной физики, том 44, № 6S. С. 4689, Июнь 2005, DOI:10.1143/JJAP.4

4. *Мур Р.К.* Радиосвязь в море // IEEE Геология, том 4, № 11. С. 42-51, Ноябрь 1967, DOI:10.1109/MSPEC.1967.5217169/MSPEC.1967.5217169

5. *Шоу Э., Аль-Шамма А.И., Уайли С.Р., Тоал Д.* Экспериментальные исследования распространения электромагнитных волн в морской воде // Микроволновая конференция, Сентябрь 2006. С. 572-575, DOI:10.1109/EUMC.2006.281456

6. *Урибе К., Гроте В.* Модель радиосвязи для подводной WSN // 3-я международная конференция по новым технологиям, мобильности и безопасности, Декабрь 2009. С. 1-5, DOI:10.1109/NTMS.2009.5384789.

7. *Дантли С.* Свет в море // Наука об окружающей среде, том 53, № 2. С. 214-233, Февраль 1963, DOI:10.1364/JOSA.53.000214

8. *Хэнсон Ф., Радич С.* Подводная оптическая связь с высокой пропускной способностью // Прикладная оптика, том. 47, № 2. С. 277-283, Январь 2008, DOI:10.1364/AO.47.000277

9. *Накамура К., Мидзукоси И., Ханава М.* Оптическая беспроводная передача 405 нм, 1,45 Гбит/с оптических сигналов IM/DD-OFDM через подводный канал длиной 4,8 м // Оптика Экспресс, том. 23, № 2. С. 1558-1566, Январь 2015, DOI:10.1364/OE.23.001558

10. *Мобли К.* Свет и вода: перенос излучения в природных водах // Академическая пресса, 1994.

11. *Джонсон Л.Д., Ясман Ф., Грин Р., Лисон М.С.* Последние достижения в области подводной оптической беспроводной связи // Подводные технологии Международный журнал подводного общества, том 32, № 3. С. 167-175, Ноябрь 2014, DOI:10.3723/ut.32.167.

12. *Санчес Р., Маккорми Н.Дж.* Аналитическая функция распространения луча для приложений океанской оптики // Прикладная оптика, том 41, № 30. С. 6276-6288, Октябрь 2002, DOI:10.1364/AO.41.006276.

13. *Яруватанадилок С.* Моделирование и оценка производительности подводного беспроводного оптического канала связи с использованием векторной теории переноса излучения // Журнал IEEE по отдельным областям связи, том 26, № 9. С. 1620-1627, Декабрь 2008, DOI:10.1109/JSAC.2008.081202.

14. *Габриель Ч., Халига М.А., Буреннан С., Леон П., Ригауд В.* Соображения о несоосности в подводных беспроводных оптических линиях связи типа "Точка-точка", Конф. OCEAN'S 2013, Июнь 2013. С. 1-5, DOI:10.1109/OCEANS-Bergen.2013.6607990.

15. *Халига М.А., Уйсал М.* Обзор оптической связи в свободном пространстве: перспектива теории связи // Обзоры и учебные пособия по коммуникациям IEEE, том 16, № 4. С. 2231-2258, Июнь 2014, DOI:10.1109/COMST.2014.2329501.

16. *Ахунди Ф., Салехи Дж.А., Ташакори А.* Сотовая подводная беспроводная оптическая сеть CDMA: анализ производительности и концепции внедрения // IEEE Операции по коммуникациям, том 63, № 3. С. 882-891, Март 2015, DOI:10.1109/TCOMM.2015.2400441.

17. *Гассемлуи З., Попула В., Раджбхандари С.* Оптическая беспроводная связь: моделирование системы и канала с помощью MATLAB(R) // CPC, Август 2012, DOI:10.1201/b12687.

18. *Шиу Д.-С., Кан Дж.М.* Дифференциальная импульсно-позиционная модуляция для энергоэффективной оптической связи // IEEE Операции по коммуникациям, том 47, № 8. С. 1201-1210, Август 1999, DOI:10.1109/26.780456.

19. *Гассемлуи З., Алдиббиат Н.М.* Многоуровневая цифровая схема импульсной интервальной модуляции для оптической беспроводной связи // Международная конференция 2006 года по прозрачным оптическим сетям, Ноттингем, Великобритания, Июнь 2006, том 3. С. 149-153, DOI:10.1109/ICTON.2006.248423.

20. *Сари Х., Вудворд Б.* Подводная голосовая связь с использованием модулированного лазерного луча // Материалы конференции Океаны '98, том 2, Сентябрь 1998. С. 1183-1188, DOI:10.1109/OCEANS.1998.724422.

21. *Чен М., Чжоу С., Ли Т.* Реализация PPM в подводной лазерной системе связи // Коммуникации, схемы и системы, 2006 Международная конференция, том 3, Июнь 2006. С. 1901-1903, DOI:10.1109/ICCCAS.2006.285044.

22. *Мейхонг С., Синьшен Ю., Фенли Ч.* Оценка методов модуляции для подводной беспроводной оптической связи // Коммуникационное программное обеспечение и сети ICCSN '09, Февраль 2009. С. 138-142, DOI:10.1109/ICCSN.2009.97.

23. *Ангита Д., Бриззолара Д., Пароди Г.* VHDL моделирование модулей уровня PHY и MAC для подводной оптической беспроводной связи // Материалы 5-й Европейской конференции по схемам и системам связи (ECCSC'10), Белград, Сербия, Ноябрь 2010. С. 185-188, DOI:10.1109/ACCESS.2022.3225913.

24. *Ангита Д., Бриззолара Д., Пароди Г.* Оптическая беспроводная связь для подводных беспроводных сенсорных сетей: проектирование и реализация аппаратных модулей и схем // OCEANS 2010 MTS/IEEE SEATTLE, Сиэтл, Вашингтон, США, 2010. С. 1-8, DOI:10.1109/OCEANS.2010.5664321.

25. *Хагем Р.М., Тиль Д.В., О'Киф С.Г., Фикеншер Т.* Оптическая беспроводная связь для обратной связи пловцов в режиме реального времени: обзор // 2012 Международный симпозиум по коммуникациям и информационным технологиям (ISCIT), Голд-Кост, Квинсленд, Австралия, 2012. С. 1080-1085, DOI:10.1109/ISCIT.2012.6380853.

26. *Хе Х., Ян Дж.* Исследование производительности подводных оптических систем связи M-ary PPM с использованием теории векторного переноса излучения // ISAPE2012, Сиань, Китай, 2012. С. 566-570, DOI:10.1109/ISAPE.2012.6408834.

27. *Свати П., Принс С.* Проблемы проектирования подводной беспроводной оптической системы связи // Международная конференция 2014 года по связи и обработке сигналов, Мелмаруватур, Индия, 2014. С. 1440-1445, DOI:10.1109/ICCSP.2014.6950087.

28. *Габриель К., Халиги М.А., Буреннан С., Леон П., Ригауд В.* Исследование подходящих методов модуляции для подводной беспроводной оптической связи // 2012 Международный семинар по оптической беспроводной связи (IWOW), Пиза, Италия, 2012. С. 1-3, DOI:10.1109/IWOW.2012.6349691.

29. *Дониец М., Василеску И., Читре М., Детвейлер К., Хоффманн-Кунт М., Рус Д.* AquaOptical: легкое устройство для высокоскоростной подводной связи "точка-точка" на большие расстояния// ОКЕАНЫ 2009, Билокси, Миссисипи, США, 2009. С. 1-6, DOI:10.23919/OCEANS.2009.5422200.

30. *Дониец М., Рус Д.* Двунаправленная оптическая связь с AquaOptical II // 2010 Международная конференция IEEE по системам связи, Сингапур, 2010. С. 390-394, DOI:10.1109/ICCS.2010.5686513.

31. *Дониец М., Детвейлер К., Василеску И., Рус Д.* Использование оптической связи для удаленной работы подводного робота // 2010 Международная конференция IEEE /RSJ по интеллектуальным роботам и системам, Тайбэй, Тайвань, 2010. С. 4017-4022, DOI:10.1109/IROS.2010.5650224.

32. *Коченор Б., Маллен Л., Лаукс А.* Фазово-когерентная цифровая связь для беспроводных оптических линий связи в мутных подводных средах // OCEANS 2007, Ванкувер, Британская Колумбия, Канада, 2007. С. 1-5, DOI:10.1109/OCEANS.2007.4449173.

33. *Кокс У.К., Хьюз Б.Л., Мут Дж.Ф.* Система управления поляризацией для подводной оптической связи // ОКЕАНЫ 2009, Билокси, Миссисипи, США, 2009. С. 1-4, DOI:10.23919/OCEANS.2009.5422258.

34. *Донг Ю., Чжан Т., Чжан Х.* Модуляция положения поляризованного импульса для беспроводной оптической связи // 47-я ежегодная конференция по информационным наукам и системам (CISS) 2013 года, Балтимор, Мэриленд, США, 2013. С. 1-5, DOI:10.1109/CISS.2013.6624253.

35. *Гальярди Р.М., Карп С.* Оптическая связь // Нью-Йорк, Уайли-Интернаука, 1976. 445 с., DOI:10.1007/1-4020-0613-6_12948.

36. *Сюй Ф., Халиги М.А., Буреннан С.* Влияние различных источников шума на производительность приемников FSO на основе PIN и APD// Материалы 11-й международной конференции по телекоммуникациям, Грац, Австрия, 2011. С. 211-218.

37. *Дэвидсон Ф.М., Сан С.* Аппроксимация по Гауссу в сравнении с почти точным анализом производительности оптических систем связи с сигнализацией PPM и приемниками APD // в IEEE Операции по коммуникациям, том 36, № 11. С. 1185-1192, Ноябрь 1988, DOI:10.1109/26.8924.

OVERVIEW OF KEY FEATURES IN THE CONSTRUCTION OF UNDERWATER OPTICAL WIRELESS COMMUNICATION

IVAN I. PAVLOV
St. Petersburg, Russia, iipavlov02@mail.ru

MARIA S.PAVLOVA
St. Petersburg, Russia, mspavlova@ngs.ru

EVGENIA S. ABRAMOVA
St. Petersburg, Russia, evgenka_252@mail.ru

SERGEY S. ABRAMOV
St. Petersburg, Russia, abramov@sibsutis.ru

YURIY S. SHCHERBAKOV
St. Petersburg, Russia, ampal55@mail.ru

KEYWORDS: *optical wireless communication channel, underwater optical wireless communication, information transfer rate, communication range, laser light beam, type of ocean water, intensity modulation, coherent modulation.*

ABSTRACT

Intoduction: increased interest in the researches of the underwater environment of the oceans is required for the study of flora and fauna, the ocean floor, the search for minerals, monitoring existing oil rigs and other objects in the ocean, collecting the necessary information. All this cannot be done without high-quality and reliable communication in underwater oceanic conditions. Known methods of organizing underwater wireless communication considered communication systems using acoustic, radio frequency and optical waves. Foreign and Russian researchers note that underwater acoustic and radio frequency wireless communication is characterized by a low speed of information transmission and high latency, but at the same time, it allows you to transmit information over long distances. The need for high-speed information transmission has contributed to the development of underwater optical wireless communication. It allows you to organize communication at a high speed of information transmission with low latency over short distances. An increase in the communication range may provided with the help of network technology for building such communication systems. Purpose: the purpose is to review the scientific works of foreign researchers in the analysis of theoretical and experimental research in the field of underwater optical wireless communication. Results: It is shown, that underwater optical wireless communication maybe organized directly between the surface station and optical nodes, or using network technology to build a network to organize a multi-level communication system from the surface station to the optical nodes using optical access points and/or optical base stations. The analysis showed that when organizing underwater optical wireless communication, the main influence on the quality and reliability of communication is directly exerted by the underwater oceanic environment. Practical relevance: of the work lies in the systematization of scientific works of foreign researchers on the problems of underwater optical wireless communication and use as a basis for future applied research.

REFERENCES

1. M. Stojanovic, “Recent advances in high-speed underwater acoustic communications,” *IEEE Journal of Oceanic Engineering*, vol. 21. No. 2, pp. 125-136, Apr. 1996, DOI:10.1109/48.486787

2. A. Zielinski, Y.H. Yoon, L. Wu, “Performance analysis of digital acoustic communication in a shallow water channel,” *IEEE J. Ocean. Eng.,* vol. 20. No. 4, pp. 293-299, Oct. 1995, DOI: 10.1109/48.468244.

3. H. Ochi, Y. Watanabe, T. Shimura, “Basic Study of Underwater Acoustic Communication Using 32-Quadrature Amplitude Modulation,” *Japanese J. of App. Physics*, vol. 44. No. 6S, p. 4689, Jun. 2005, DOI:10.1143/JJAP.4.

4. R.K. Moore, “Radio communication in the sea,” *IEEE Spectrum*, vol. 4. No. 11, pp. 42-51, Nov. 1967, DOI:10.1109/MSPEC.1967.5217169/MSPEC.1967.5217169.

5. A. Shaw, A.I. Al-Shamma'a, S.R. Wylie, D. Toal, “Experimental investigations of electromagnetic wave propagation in seawater,” *European Microwave Conf.*, Sep. 2006, pp. 572-575, DOI:10.1109/EUMC.2006.281456.

6. C. Uribe, W. Grote, “Radio Communication Model for Underwater WSN,” *3rd Int. Conf. on New Technologies, Mobility and Security*, Dec. 2009, pp. 1-5, DOI:10.1109/NTMS.2009.5384789.

7. S.Q. Duntley, “Light in the sea,” *J. Opt. Soc. Am*., vol. 53. No. 2, pp. 214-233, Feb. 1963, DOI:10.1364/JOSA.53.000214.

8. F. Hanson, S. Radic, “High bandwidth underwater optical communication,” *Appl. Opt*., vol. 47. No. 2, pp. 277-283, Jan. 2008, DOI:10.1364/AO.47.000277.

9. K. Nakamura, I. Mizukoshi, M. Hanawa, “Optical wireless transmission of 405 nm, 1.45 Gbit/s optical IM/DD-OFDM signals through a 4.8 m underwater channel,” *Opt. Express*, vol. 23. No. 2, pp. 1558-1566, Jan. 2015, DOI:10.1364/OE.23.001558

10. C. Mobley, “Light and Water: Radiative Transfer in Natural Waters,” Academic Press, 1994.

11. L.J. Johnson, F. Jasman, R. Green, M.S. Leeson, “Recent advances in underwater optical wireless communications,” *Underwater Techno*., vol. 32. No. 3, pp. 167-175, Nov. 2014, DOI:10.3723/ut.32.167.

12. R. Sanchez, N.J. McCormick, “Analytic beam spread function for ocean optics applications,” *Appl. Opt.*, vol. 41. No. 30, pp. 6276-6288, Oct. 2002, DOI:10.1364/AO.41.006276.

13. S. Jaruwatanadilok, “Underwater wireless optical communication channel modeling and performance evaluation using vector radiative transfer theory,” *IEEE J. Sel. Areas Commun*., vol. 26. No. 9, pp. 1620-1627, Dec. 2008, DOI:10.1109/JSAC.2008.081202.

14. C. Gabriel, M.A. Khalighi, S. Bourennane, P. Lon, V. Rigaud, “Misalignment considerations in point-to-point underwater wireless optical links,” *OCEANS*, Jun. 2013, pp. 1-5, DOI:10.1109/OCEANS-Bergen.2013.6607990.

15. M.A. Khalighi, M. Uysal, “Survey on Free Space Optical Communication: A Communication Theory Perspective,” *IEEE Commun. Surveys Tuts.,* vol. 16. No. 4, pp. 2231-2258, Jun. 2014, DOI:10.1109/COMST.2014.2329501.

16. F. Akhoundi, J.A. Salehi, A. Tashakori, “Cellular Underwater Wireless Optical CDMA Network: Performance Analysis and Implementation Concepts,” *IEEE Trans. Commun.*, vol. 63. No. 3, pp. 882-891, Mar. 2015, DOI:10.1109/TCOMM.2015.2400441.

17. Z. Ghassemlooy, W. Popoola, S. Rajbhandari, “Optical Wireless Communications: System and Channel Modelling with MATLAB(R),” *CRC*, Aug 2012, DOI:10.1201/b12687.

18. D.-S. Shiu, J.M. Kahn, “Differential pulse-position modulation for power-efficient optical communication,” *IEEE Trans. Commun*., vol. 47. No. 8, pp. 1201-1210, Aug. 1999, DOI:10.1109/26.780456.

19. Z. Ghassemlooy, N.M. Aldibbiat, “Multilevel digital pulse interval modulation scheme for optical wireless communications,” *Int. Conf. on Transparent Opt. Netw.,* vol. 3, pp. 149-153, Jun. 2006, DOI:10.1109/ICTON.2006.248423.

20. H. Sari, B. Woodward, “Underwater voice communications using a modulated laser beam,” *OCEANS '98 Conference Proceedings*, vol. 2, Sep. 1998, pp. 1183-1188, DOI:10.1109/OCEANS.1998.724422.

21. M. Chen, S. Zhou, T. Li, “The Implementation of PPM in Underwater Laser Communication System,” *Communications, Circuits and Systems Proceedings, 2006 International*, vol. 3, Jun. 2006, pp. 1901-1903, DOI:10.1109/ICCCAS.2006.285044.

22. S. Meihong, Y. Xinsheng, Z. Fengli, “The evaluation of modulation techniques for underwater wireless optical communications,” *Communication Software and Networks*, ICCSN '09, Feb. 2009, pp. 138-142, DOI:10.1109/ICCSN.2009.97.

23. D. Anguita , D. Brizzolara, G. Parodi, “VHDL modeling of PHY and MAC Layer modules for underwater optical wireless communication,” *Proceedings of Papers 5th European Conference on Circuits and Systems for Communications (ECCSC'10)*, Belgrade, Serbia, Nov. 2010, pp. 185-188, DOI:10.1109/ACCESS.2022.3225913.

24. D. Anguita, D. Brizzolara, G. Parodi, “Optical wireless communication for underwater Wireless Sensor Networks: Hardware modules and circuits design and implementation,” *OCEANS 2010 MTS/IEEE SEATTLE*, Seattle, WA, USA, 2010, pp. 1-8, DOI:10.1109/OCEANS.2010.5664321.

25. R.M. Hagem, D.V. Thiel, S.G. O'Keefe, T. Fickenscher, “Optical wireless communication for real time swimmers feedback: A review,” *2012 International Symposium on Communications and Information Technologies (ISCIT)*, Gold Coast, QLD, Australia, 2012, pp. 1080-1085, DOI:10.1109/ISCIT.2012.6380853.

26. X. He, J. Yan, “Study on performance of M-ary PPM underwater optical communication Systems using vector radiative transfer theory,” *ISAPE2012*, Xi'an, China, 2012, pp. 566-570, DOI:10.1109/ISAPE.2012.6408834.

27. P. Swathi, S. Prince, “Designing issues in design of underwater wireless optical communication system,” *2014 International Conference on Communication and Signal Processing*, Melmaruvathur, India, 2014, pp. 1440-1445, DOI:10.1109/ICCSP.2014.6950087.

28. C. Gabriel, M.A. Khalighi, S. Bourennane, P. Lon, V. Rigaud, “Investigation of suitable modulation techniques for underwater wireless optical communication,” *2012 International Workshop on Optical Wireless Communications (IWOW)*, Pisa, Italy, 2012, pp. 1-3, DOI:10.1109/IWOW.2012.6349691.

29. M. Doniec, I. Vasilescu, M. Chitre, C. Detweiler, M. Hoffmann-Kuhnt, D. Rus, “Aquaoptical: A lightweight device for high-rate long-grange underwater point-to-point communication,” *OCEANS 2009,* Biloxi, MS, USA, 2009, pp. 1-6, DOI:10.23919/OCEANS.2009.5422200.

30. M. Doniec, D. Rus, “BiDirectional optical communication with AquaOptical II,” *2010 IEEE International Conference on Communication Systems, Singapore*, 2010, pp. 390-394, DOI:10.1109/ICCS.2010.5686513.

31. M. Doniec, C. Detweiler, I. Vasilescu, D. Rus, “Using optical communication for remote underwater robot operation,” *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems,* Taipei, Taiwan, 2010, pp. 4017-4022, DOI:10.1109/IROS.2010.5650224.

32. B. Cochenour, L. Mullen, A. Laux, “Phase coherent digital communications for wireless optical links in turbid underwater environments,” *OCEANS 2007*, Vancouver, BC, Canada, 2007, pp. 1-5, DOI:10.1109/OCEANS.2007.4449173.

33. W.C. Cox, B.L. Hughes, J.F. Muth, “A polarization shiftkeying system for underwater optical communications,” *OCEANS 2009*, Biloxi, MS, USA, 2009, pp. 1-4, DOI:10.23919/OCEANS.2009.5422258.

34. Y. Dong, T. Zhang, X. Zhang, “Polarized pulse position modulation for wireless optical communications,” *2013 47th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, 2013, pp. 1-5, DOI:10.1109/CISS.2013.6624253.

35. R.M. Gagliardi, S. Karp, “Optical Communications,” *New York, Wiley-Interscience*, 1976. 445 p., DOI:10.1007/1-4020-0613-6_12948.

36. F. Xu, M.A. Khalighi, S. Bourennane, “Impact of different noise sources on the performance of PIN- and APD-based FSO receivers,” *Proceedings of the 11th International Conference on Telecommunications*, Graz, Austria, 2011, pp. 211-218.

37. F.M. Davidson, X. Sun, “Gaussian approximation versus nearly exact performance analysis of optical communication systems with ppm signaling and apd receivers,” *IEEE Transactions on Communications*, vol. 36. No. 11, pp. 1185-1192, Nov. 1988, DOI:10.1109/26.8924.

INFORMATION ABOUT AUTHORS:

Ivan I.Pavlov, PhD, docent, associate Professor at the Department of "Radio engineering devices and technosphere security", federal state institution of higher education "Siberian state university of telecommunication and information science", Novosibirsk, Russia, iipavlov02@mail.ru

Maria S.Pavlova, PhD, associate Professor at the Department of "Radio engineering devices and technosphere security", federal state institution of higher education "Siberian state university of telecommunication and information science", Novosibirsk, Russia, mspavlova@ngs.ru

Evgenia S.Abramova, PhD, docent, associate Professor at the Department of "Radio engineering devices and technosphere security", federal state institution of higher education "Siberian state university of telecommunication and information science", Novosibirsk, Russia, evgenka_252@mail.ru

Sergey S.Abramov, PhD, docent, head of the department of "Radio engineering devices and technosphere security", federal state institution of higher education "Siberian state university of telecommunication and information science", Novosibirsk, Russia, abramov@sibsutis.ru

Yuriy S.Shcherbakov, PhD, docent, associate Professor at the Department of "Radio engineering devices and technosphere security", federal state institution of higher education "Siberian state university of telecommunication and information science", Novosibirsk, Russia, ampal55@mail.ru

For citation: Pavlov I.I, Pavlova M.S., Abramova E.S., Shcherbakov Yu.S., Shcherbakov Yu.S. Overview of key features in the construction of underwater optical wireless communication. H&ES Reserch. 2023. Vol. 15. No. 4. P. 14-25. doi: 10.36724/2409-5419-2023-15-4-14-25 (In Rus)

doi: 10.36724/2409-5419-2023-15-4-26-37

ОЦЕНКА ЭЛЕКТРОМАГНИТНОЙ ДОСТУПНОСТИ ИСТОЧНИКОВ РАДИОИЗЛУЧЕНИЙ ГСС STARLINK

СЕВИДОВ
Владимир Витальевич [1]

ДВОРНИКОВ
Сергей Сергеевич [2]

БЕСТУГИН
Александр Роальдович [3]

КИРШИНА
Ирина Анатольевна [4]

ДВОРНИКОВ
Сергей Викторович [5]

Сведения об авторах:

[1] к.т.н., доцент, докторант, Военная академия связи им. С.М. Буденного, Санкт-Петербург, Россия, v-v-sevidov@mail.ru
orcid.org/0009-0009-2413-1615

[2] к.т.н., доцент, кафедра конструирования и технологий электронных и лазерных средств, Санкт-Петербургский государственный университет аэрокосмического приборостроения; научный сотрудник, научно-исследовательский отдел, Военная академия связи им. С.М. Буденного, Санкт-Петербург, Россия, dvornik.92@mail.ru
orcid.org/0000-0001-7426-6475

[3] д.т.н., профессор, директор интитута радиотехники и инфокоммуникационных технологий, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия, fresguap@mail.ru,
orcid.org/0000-0003-3847-2516

[4] к.э.н., доцент, кафедра конструирования и технологий электронных и лазерных средств, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия, ikirshina@mail.ru,
orcid.org/0000-0002-31102363

[5] д.т.н., профессор, кафедра радиотехнических и оптоэлектронных комплексов, Санкт-Петербургский государственный университет аэрокосмического приборостроения; профессор кафедры радиосвязи, Военная академия связи им. С.М. Буденного, Санкт-Петербург, Россия, practicdsv@yandex.ru
orcid.org/0000-0001-7426-6475

АННОТАЦИЯ

Введение. Полномасштабное развертывание глобальной спутниковой системы Starlink открыло широкие возможности по высокоскоростному беспроводному доступу к информационным ресурсам сети Internet на всей территории Земли. В таких условиях возникает проблема контроля пользователей за предоставляемым трафиком. Это обусловлено тем, что организационно технические параметры системы Starlink изначально выбирались с условием минимизации взаимных помех абонентам, находящимся на ограниченных территориях. С одной стороны такая система открывает перспективу доступа к глобальной сети абонентам, находящимся в удаленных, индустриально необорудованных районах Земного шара, а с другой – возможность организации связи для нелегитимных пользователей и право преступных организаций. Поэтому необходимо понимание того, насколько доступны для мониторинга каналы связи, организуемые системой Starlink. **Цель исследования:** оценка вероятности электромагнитной доступности абонентов глобальной спутниковой системы Starlink. **Методы:** На основе проведенного анализа в доступных источниках оперативных и технических характеристик системы Starlink, были рассмотрены особенности ее построения и определены исходные данные для разработки методики электромагнитной доступности источников радиоизлучений, работающих в Ku/K-диапазонах частот. В основе методики лежат расчеты обеспечиваемого превышения мощности сигнала над мощностью шума на входе радиоприемного устройства, при которых обеспечивается требуемое качество и достоверность приема. Обоснованы этапы методики, модифицирован аналитический аппарат с учетом условий проведения мониторинга. **Представлены результаты** расчета электромагнитной доступности источников радиоизлучений глобальной спутниковой системы "Starlink для наиболее сложного случая, когда абонентский терминал находится в надире зоны покрытия космическим аппаратом и одновременно в центре луча, формируемого на абонента. Получены вероятностные оценки и сформулированы предложения и рекомендации по повышению надежности и качества проводимого маниторинга. Сформулированы выводы и предложения по практической реализации полученных результатов.

КЛЮЧЕВЫЕ СЛОВА: *радиомониторинг абонентов системы Starlink, электромагнитная доступность источников радиоизлучений, методика оценки электромагнитной доступности абонентов в Ku/K-диапазонах частот, помехоустойчивость приема сигналов системы Starlink.*

Для цитирования: Севидов В.В., Дворников С.С., Бестугин А.Р., Киршина И.А., Дворников С.В. Оценка электромагнитной доступности источников радиоизлучений ГСС Starlink // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 26-37. doi: 10.36724/2409-5419-2023-15-4-26-37

Введение

В 2015 году компания SpaceX приступила к разработке глобальной спутниковой системы (ГСС) Starlink, предназначенной для обеспечения высокоскоростным широкополосным спутниковым доступом в сети Internet на всей территории Земли [1, 2]. И уже через 3 года успешно система успешно прошла тестовые испытания [3, 4].

Началом работы ГСС Starlink принято считать начало мая 2019 года, когда на орбиту была выведена первая группировка из 60 спутников. И с 2020 года компания SpaceX начала предоставлять коммерческие услуги доступа к сети Internet в труднодоступных северных районах на территориях США и Канады [10].

По состоянию на январь 2022 года число пользователей бета-тестеров ГСС Starlink достигло 145 тыс. в 25 странах мира, а буквально через четыре месяца их стало более 400 тысяч [1, 5].

Космические аппараты ГСС Starlink имеют относительно небольшие размеры (порядка 260 кг) и выполнены в виде плоской панели, что обеспечивает возможность размещать на их поверхности четыре фазированные антенные решетки (ФАР) [6]. При этом спутники изготовлены на основе, так называемых «зеленых технологий», они оснащены электростатическими двигателями, работающими на эффекте Холла с использованием криптона [7, 8]. Это позволяет корректировать их орбиту в процессе эксплуатации и в конце срока службы выводить их в плотные слои атмосферы для утилизации. К концу 2020 года компания SpaceX получила выгодный контракт на предоставление доступа к высокоскоростному Internet ресурсу пользователя на тех территориях США, где он раньше был не доступен [9].

Возможно ГСС Starlink так и осталась бы специфическим амбициозным проектом Илона Маска, если бы им не заинтересовалось министерство обороны США. Еще в декабре 2018 года ВВС США заключили со компанией SpaceX серьезный контракт, предусматривающий изучение различных вариантов использования ГСС Starlink для военных нужд. В 2019 году, в рамках тестовых испытаний производительности сервиса, была достигнута пропускная способность 610 Мбит/с терминала, размещенного на борту военно-транспортного самолёта C-12J Huron. После чего министерство обороны США подписала с компанией SpaceX соглашение, рассчитанное на три года, по тестированию системы передачи данных между своими наземными подразделениями посредством использования космического сегмента [1]. К середине сентябре 2020 года компания SpaceX достигла планируемой скорости в 100 Мбит/с при сверхнизкой задержке сигнала, а также возможность использования лазера между двумя спутниками на орбите [10].

После этого специалистами пентагона было сделано заключение о потенциальной возможности использования низкоорбитальных спутниковых систем типа ГСС Starlink в качестве надежной альтернативы навигационной системы GPS [11]. В августе 2022 года, Командование ВВС США в Европе и Африке подписало контракт по использованию услуг ГСС Starlink в интересах поддержки 86-го авиакрыла, размещенного на авиабазе Рамштайн в Германии, в том числе, и в интересах группировки войск Украины в ходе проведения специальной военной операции. При этом, в обосновывающем контракт документе, ГСС Starlink позиционируется как единственная коммерческая система, способная предоставлять широкий спектр услуг на всей территории Земли [1].

Указанные обстоятельства стимулировали к проведению исследований оценки вероятности электромагнитной доступности (ЭМД) абонентов ГСС «Starlink», результаты которой представлены в настоящей статье.

Анализ особенностей организационно-технического построения ГСС Starlink

Первоначально ГСС Starlink предусматривала размещение группировки космических аппаратов (КА) на высоте порядка 1100-1325 км над поверхностью Земли с наклонениями орбит 53,8°, 70°, 74° и 81°. Но затем в ходе тестирования было принято решение о снижении высот до уровня 540-570 км, с наклоном орбит 53,2°, 70° и 97,6° и общим количеством спутников в 4408 единиц [12, 13].

В рамках предоставляемых услуг связи ГСС Starlink обеспечивает скорость доступа в Internet до 500 Мбит/с без ограничения объема трафика, но при отсутствии других абонентов в ячейке, формируемой посредством луча ФАР, размещенной на борту КА, и задержкой предоставления трафика в пределах 20-40 мс [14].

Когда к июню 2022 года число пользователей ГСС Starlink превысило 500 000 абонентских номеров, скорость предоставления услуг в пакете «Starlink Business» снизилась до 150-500 Мбит/с [15]. И уже в августе 2022 года было объявлено о совместном с компанией T-Mobile проекте по предоставлению услуг мобильной связи на базе ГСС Starlink. В указанном соглашении предусматривалось оснащение КА Starlink дополнительным комплектом оборудования PCS (personal communications service), обеспечивающего 1000-2000 телефонных вызовов и порядка 1 млн коротких текстовых сообщений в зоне покрытия со скоростью от 2 до 4 Мбит/с, и возможностью совместной работы со стандартными сотовыми телефонами [1].

Основной полезной нагрузкой КА Starlink являются два антенных комплекса для связи со шлюзовыми станциями (ШС) (гейтвеями) и с абонентскими терминалами (АТ) [16]. Антенный комплекс для связи с ШС работает в K-диапазоне (18/30 ГГц) [6]. Характеристики каналов связи ГСС Starlink представлены в таблице 1.

Согласно данным таблицы 1, частотный ресурс радиолинии, организуемой от шлюзовой станции к КА, составляет 2100 МГц, а в обратном направлении – 1300 МГц. В случае использования двух вариантов поляризации (левой и правой, или в случае применения антенн с эллиптической поляризацией) частотный ресурс удваивается.

При этом непосредственная связь между АТ и КА организуется в Ku-диапазоне в полосе 2 ГГц на радиолинии вниз от КА к АТ и порядка 0,5 ГГц на радиолинии вверх (данные лишь для одного вида поляризации, при использовании обоих видов поляризации частотный ресурс удваивается).

Таблица 1

Характеристики каналов связи, организуемых посредством ГСС Starlink

Тип канала связи и направление приема-передачи	Диапазоны частот, ГГц	Доступно МГц в одной поляризации
Услуга: нисходящий канала (КА – АТ)	10,7-12,7	2000
Услуга: нисходящий канала (КА – ШС)	17,8-18,6 18,8-19,3	800 500
Услуга: восходящий канала (АТ – КА)	14,0-14,5	500
Услуга: восходящий канала (ШС – КА)	27,5-29,1 29,5-30,0	1600 500
Телеметрия и управление вниз (нисходящий канал КА - станция контроля)	12,15-12,25 18,55 - 18,60	100 50
Телеметрия и управление вверх (восходящий канала станция контроля - КА)	13,85 - 14,00	150

На борту КА размещен комплект оборудования для организации работы командной радиолинии и передачи данных телеметрии в К и Ku диапазонах с полосой канала 150 МГц [5].

Работа телекоммуникационного оборудования, размещенного на борту КА Starlink, осуществляется в режиме ретранслятора, то есть без обработки информации. На его борту происходит лишь конвертирование частоты и дополнительное усиление сигнала.

Находясь на орбите высотой порядка 550 км, каждый КА посредством своих антенных систем формирует зону покрытия на поверхности Земле радиусом около 950 км, с учетом того, что угол места для АТ будет не менее 25°. При этом эффективная работа антенн с плоской фазированной решеткой возможна при угле места 40° и более [1]. В табл. 2 приведены основные технические характеристики радиолиний, существенные для организации электромагнитной доступности абонентов.

Таблица 2

Основные технические характеристики радиолиний, организуемых посредством ГСС Starlink

Тип канала	Частота, ГГц	Вид модуляции	Максимальная ЭИИМ	Ширина луча ДН половинной мощности
Нисходящий канал (КА – АТ)	10,7-12,7	от QAM-2 до QAM-64	Нет данных	3,5° (boresight) 5,5° (at slant)
Восходящий канал (АТ – КА)	14,0-14,5	от QAM-2 до QAM-64	38,2 дБВт	2,8° (boresight) 4,5° (at slant)

В пределах зоны покрытия одного КА формируется не более 16 лучей, между которыми распределяется выделяемый частотный ресурс Ku-диапазона. Общий подход показан на рисунке 1.

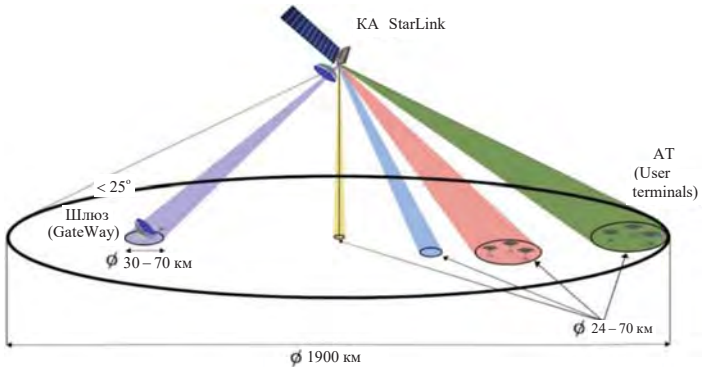


Рис. 1. Принцип формирования лучей в пределах зоны покрытия КА ГСС Starlink при организации работы радиолиний вниз

В соответствии с конфигурацией, представленной на рис. 1, общая площадь зоны ЭМД КА ГСС Starlink с углом места до 25° и высоте орбиты 550 км достигает 2 835 294 км^2 [12]. А время функционирования радиолинии наземного АТ с КА составляет всего 4,1 мин или около 250 с. При этом, в соответствии с топологией зоны ЭМД, наибольшая продолжительность непрерывной работы радиолинии КА – АТ обеспечивается при нахождении наземного абонентов не в надире (подспутниковой точке), а в периферийной зоне видимости, по направлению пролета КА.

Доступ наземного АТ к сети Internet обеспечивается через ШС. При этом одна ШС способна одновременно распределять Internet-трафик тысячам терминалов. Типовая ШС имеет 8 параболических антенн (ПА) с диаметром зеркала 1,5 м, подключенных к передатчикам мощностью 50 Вт [13].

Применение параболических антенн в отличие от фазированных решеток позволяет работать даже при углах места до 5° [1, 6]. При этом терминал компании SpaceX в канале 500 МГц (с учетом защитных интервалов, до 480 МГц) обеспечивает эквивалентную изотропную излучаемую мощность (ЭИИМ) до 66,5 дБВт.

Эксплуатационные возможности АТ ГСС Starlink гораздо скромнее. Первоначально компанией SpaceX было заявлено пять типов АТ, которые позиционируются как Model "ES A (B, C, D. E)" и два типа терминалов ШС (см. табл. 3 и 4) [1].

Таблица 3

Основные технические характеристики терминалов шлюзовых станций

Основные Элементы	Model “Space* Telem Ku/K"	Model “Space* Telem X"
Тип модели	CGC Technology T450	Orbital Systems 3.7 Meter
Тип антенны	ПА Кассегрена	ПА с основным фокусом
Диапазон частот	Работа во всем диапазоне заявленных частот	2,0-2,1 ГГц восходящий канал 7,2-8,4 ГГц нисходящий канал
Максимальное усиление / Ширина ДН по уровню спада 3 дБ	Восходящий канал (Ku): 26 дБ 0,22°. Нисходящая линия (Ku): 22 дБ 0,3°. Нисходящий канал (K): 27 дБ 0,2°.	Восходящий канал: 8,9 дБ; 1,7. Нисходящий канал: 16,8 дБ; 0,8°

Терминалы ШС предназначены для решения задач управления и мониторинга за КА.

Таблица 4

Основные технические характеристики абонентских терминалов

Элементы	Model "ES A"	Model "ES-B"	Model "ES-C"	Model "ES-D"	Model "ES- E"
Тип модели	SpaceX	SpaceX	SpaceX	SpaceX	CGC Technology T450
Тип антенны	ФАР	ФАР	Нет данных	ПА	ПА Кас-сегрена
Диапазон частот	Работа во всем диапазоне заявленных частот	Работа во всем диапазоне заявленных частот	Работа во всем диапазоне заявленных частот	Работа во всем диапазоне заявленных частот	Работа во всем диапазоне заявленных частот
Максимальное усиление / Ширина ДН по уровню спада 3 дБ	3 дБ при полном луче 3,5°	7 дБ при полном луче 2°	6 дБ при полном луче 2,4°	16 дБ при полном луче 0,9°	Восходящий канал: 26 дБ; 0,22°. Нисходящий канал: 22 дБ; 0,3°

При этом все представленные АТ обеспечивают работу с каналами на прием (нисходящий канал) в диапазоне частот 15, 30, 60, 120 и 240 МГц с символьной скоростью 15, 30, 60, 120 и 240 МБод/с. Информационная скорость входящего потока достигает 350 Мбит/с.

Скорость Internet трафика достигает 130 Мбит/с. Как правило, работа АТ осуществляется в полосе 60 МГц. И при эффективном диаметре антенны 48 см и угле ДН 2,8°, обеспечивается ЭИИМ 38,2 дБВт, см. табл. 2. Излучаемая мощность АТ может меняться в зависимости от его наклона относительно линии в зенит. Так, для случая, когда луч антенны направлен в зенит, мощность, выдаваемая на антенну, составляет 0,76 Вт (при предельном отклонении от вертикали 4,06 Вт). Спектральная эффективность, обеспечиваемая АТ на прием при ширине канала 240 МГц и скорости передачи 350 Мбит/с, составляет всего 1,5 бит/Герц.

Анализ особенностей организационно-технического построения ГСС Starlink позволяет сделать следующее заключение, существенные для обеспечения ЭМД АТ и КА.

Основной вклад в ЭИИМ как восходящих, так и нисходящих линий радиосвязи обеспечивается за счет использования антенных систем, в основе которых лежат ФАР и ПА с высоким значением коэффициента усиления и узкими ДН.

Достаточно низкие орбиты КА приводят к кратковременному (всего 250 с) интервалу, в пределах которого возможен сеанс связи, при условии сканирующего сопровождения АТ и КА посредством управления формируемыми ДН.

Использование Ku-диапазона и K-диапазона фактически исключают ЭМД АТ при размещении средств контроля на поверхности Земли ввиду высокого уровня затухания сигнала в этих диапазонах частот.

Применение ФАР с использованием сканирующих ДН существенно ограничит возможности ЭМД средств контроля, при размещении на малоподвижных подвесных платформах.

Обоснование методики оценки вероятности ЭМД источников радиоизлучений ГСС Starlink

Термином ЭМД широко используется среди специалистов, занимающихся вопросами радиомониторинга источников радиоизлучений (ИРИ) [17-19]. Но поскольку четкое понятие ЭМД не определено ни в одном из руководящих документов, то в настоящей статье под ЭМД будем понимать вероятностную оценку возможности обнаружения, технического анализа и распознавания сигналов заданных ИРИ в интересах приема передаваемой им информации или определения направления на них с требуемой достоверностью и надежностью.

Далее будем полагать, что ЭМД обеспечивается при условии необходимого превышение мощности сигнала $P_с$ над мощностью шума $P_ш$ на входе радиоприемного устройства (РПУ) [21, 22]

$$z = \frac{P_с}{P_ш} \geq z_{тр}, \qquad (1)$$

где $z_{тр}$ – минимальная величина отношения сигнал/шум (ОСШ) по мощности, соответствующая требуемой для данных условий и вида приема достоверности и надежности.

Вместе с тем текущее значение ОСШ на входе РПУ может существенно изменяться с течением времени случайным образом ввиду различных факторов, начиная от нарушения теплового баланса в аппаратуре и заканчивая изменением сигнально-помеховой обстановки на трассе распространения радиоволн. В качестве примера на рисунке 2 представлен фрагмент входной реализации $z(t)$ с изменяющимся значением ОСШ.

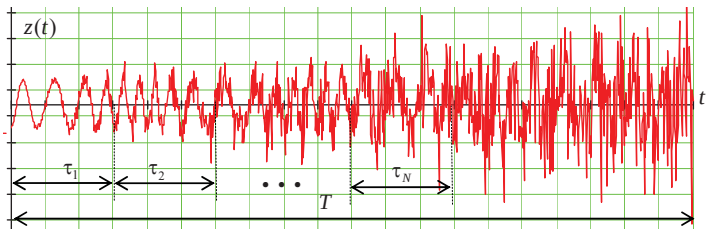


Рис. 2. Фрагмент входной реализации с изменяющимся значением ОСШ

Следовательно, о том, что будет обеспечено условие (1), т.е. что будет обеспечена ЭМД можно говорить лишь с некоторой вероятностью. Поэтому вероятность электромагнитной доступности $P_{эмд}$ целесообразно определить как вероятность того, что при общем времени наблюдения T, на N временных интервалах фрагментов длительностью τ, условие (1) выполняется [22, 23]

$$P_{эмд} = P\left(\frac{P_с}{P_ш} \geq z_{тр}\right) = \frac{\sum_N \tau_i}{T}. \qquad (2)$$

Для аналитических расчетов вероятности ЭМД более удобно использовать выражение

$$P_{\text{эмд}} = P\left(z \geq z_{\text{тр}}\right) = \int_{z_{\text{тр}}}^{\infty} W(z)\,\mathrm{d}z\,. \qquad (3)$$

где $W(z)$ – плотность распределения вероятностей измеренных значений z на входе приемника.

При большом числе измерений распределение уровней сигналов и шума во времени может быть описано нормальным законом, что позволяет условие (3) переписать к виду

$$P_{\text{эмд}} = P\left(z \geq z_{\text{тр}}\right) = \frac{1}{\sigma_z\sqrt{2\pi}} \int_{z_{\text{тр}}}^{\infty} \exp\left[-\frac{\left(\overline{z} - z_{\text{тр}}\right)^2}{2\sigma_z^2}\right]\mathrm{d}z = \Phi(u) \qquad (4)$$

где $\overline{z} = \overline{P}_{\text{с}} - \overline{P}_{\text{ш}}$ – превышение среднего уровня сигнала над средним уровнем шума; $\Phi(u)$ – интегральная функция нормального распределения случайной величины z, $u = \dfrac{\overline{z} - z_{\text{тр}}}{\sigma_z}$ – аргумент интеграла вероятности; σ_z – СКО превышения уровня сигнала над уровнем шума от его среднего значения.

С учетом сделанных допущений и веденных понятий, методику оценки вероятности ЭМД ИРИ определим совокупностью следующих основных этапов:

Этап 1. Расчет $\overline{P}_{\text{с}}$ – средней мощности сигнала на входе РПУ.

Этап 2. Расчет $\overline{P}_{\text{ш}}$ – средней мощности шума на входе РПУ.

Этап 3. Вычисление $\overline{z}$ – среднего отношения сигнал/шум.

Этап 4. Вычисление σ_z – среднеквадратического отклонения превышения уровня сигнала над уровнем шума от его среднего значения.

Этап 5. Определение $z_{\text{тр}}$ – требуемого превышения мощности сигнала над мощностью шума.

Этап 6. Вычисление u – аргумента интеграла вероятности (функции Лапласа).

Этап 7. Расчет $P_{\text{ЭМД}}$ – вероятности ЭМД.

Следует отметить, что общий вид методики оценки вероятности ЭМД ИРИ, определяемый представленными этапами 1-7, известен. Но в таком виде она имеет лишь теоретическое значение. Для практических приложений необходимо конкретизировать содержание каждого из этапов.

Поэтому далее рассмотрим особенности практического применения методики оценки вероятности ЭМД ИРИ.

Реализация первого этапа методики

В ходе передачи сигнала в радиолинии он претерпевает как усиление, так и затухание. На рисунке 3 представлена диаграмма уровней сигнала на линии связи от передатчика до приемника.

С учетом обобщения различных факторов, влияющих на среднюю мощность сигнала на входе РПУ, ее значение можно описать следующим выражением [24-26]

$$\overline{P}_{\text{с}} = P_1\eta_{\text{ф1}}G_1W_0^2W_{\text{т}}^2G_2\eta_{\text{ф2}}\xi_{\text{с}}\xi_{\text{п}}\xi_{\text{пр}}\,, \qquad (5)$$

где: P_1 – мощность передатчика ТТХ ИРИ;

$\eta_{\text{ф1}}$ и $\eta_{\text{ф2}}$ – коэффициенты полезного действия фидеров передающей и приёмной антенн;

G_1 и G_2 – коэффициенты усиления передающей и приёмной антенн (табличное значение);

$\xi_{\text{с}}$ – коэффициент согласования по сопротивлению;

$\xi_{\text{пр}}$ – коэффициент согласования по пространственной ориентации приемной и передающей антенн;

$\xi_{\text{п}}$ – коэффициент согласования по поляризации;

W_0 – множитель ослабления мощности сигнала в свободном пространстве;

$W_{\text{т}}$ – множитель ослабления сигнала на радиотрассе.

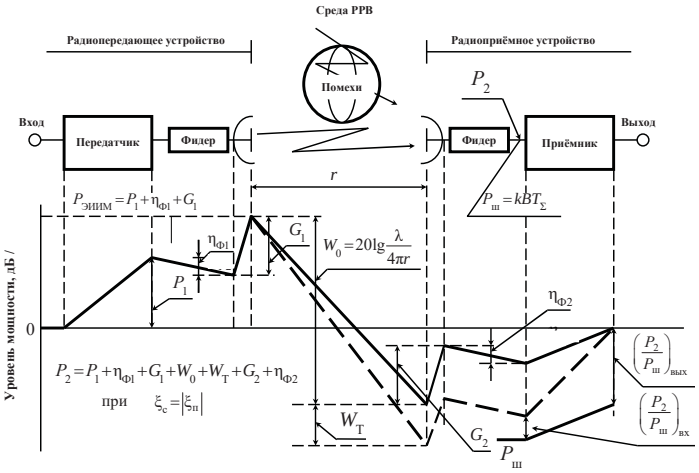


Рис. 3. Диаграмма изменения уровней сигнала в радиолинии, используемая для расчета вероятности электромагнитной доступности

Указанные значения определяют величину ЭИИМ

$$\text{ЭИИМ} = P_1 + \eta_{\text{Ф1}} + G_1\,, \qquad (6)$$

Коэффициент усиления приемной антенны G_2, в свою очередь, является функцией диаметра антенны $D_{\text{А}}$, а также частоты сигнала f и коэффициента использования поверхности (КИП) раскрыва антенны $k_{\text{А}}$, характеризующего отношение реального усиления антенны к максимально достижимому значению [27]:

$$G_2 = k_{\text{А}}\left(\frac{\pi D_{\text{А}} f}{c}\right)^2. \qquad (7)$$

Учитывая, что реальный фронт волны не является плоским, а также с учетом неравномерного распределения амплитуды поля на раскрыве ПА, величину КИП определим равной $k_{\text{А}} = 0{,}5\ldots0{,}7$.

Значение коэффициента согласования по сопротивлению $\xi_{\text{с}}$ рассчитывается по формуле:

$$\xi_{\text{с}} = \frac{4R_{\text{А}}R_{\text{Н}}}{(R_{\text{А}} + R_{\text{Н}})^2 + X_A^2}\,, \qquad (8)$$

где: $R_{\text{Н}}$ – реактивная составляющая волнового сопротивления антенны; $R_{\text{А}}$ – активная составляющая волнового сопротивления антенны.

Для увеличения значения ξ_c при работе в широком диапазоне частот, в антенных системах, как правило, используют согласующие устройства.

Величина коэффициента согласования по пространственной ориентации антенн $\xi_{пр}$ определяется взаимной ориентацией ДН приемной антенны, по отношению к передающей, и рассчитывается по формуле

$$\xi_{пр} = f_1^2(\beta_1, \theta_1)\, f_2^2(\beta_2, \theta_2), \qquad (9)$$

где (β_1, θ_1) и (β_2, θ_2) – углы между направлением максимума ДН передающей (приемной) антенны и прямой, соединяющей передающую и приемную антенну.

Значение коэффициента согласования по поляризации $\xi_п$ вводят для учета дополнительных потерь, возникающих при несовпадении поляризации антенны с принимаемой волной электромагнитного поля. Необходимо понимать, что даже в том случае, когда номинально поляризация передающей и приемной антенн совпадают, коэффициент $\xi_п$, как правило, оказывается меньшим единицы. Это связано с тем, что при распространении радиоволн в большинстве случаев происходит поворот вектора поляризации случайным образом. Поэтому, при решении большинства практических задач, связанных с расчетом вероятности ЭМД, коэффициент согласования по поляризации для каналов спутниковой связи принимают равным $\xi_п = 0{,}5\ldots1$.

Множитель ослабления мощности сигнала в свободном пространстве рассчитывается согласно формуле [19, 28]

$$W_0 = \frac{\lambda}{4\pi r}, \qquad (10)$$

и зависит не только от расстояния r между ИРИ и точкой приема, но и от длины волны λ (частоты f) излучения. Электромагнитное поле более высоких частот испытывает большее затухание в свободном пространстве.

Множитель ослабления сигнала на радиотрассе $W_т$, прежде всего, зависит от механизма их распространения, наличия препятствий, а также величины их рассеивания, определяемой параметрами среды.

Для расчета множителей ослабления на трассе $W_т$ в децибелах используют формулы, предложенные в рекомендациях МСЭ-R «Процедура прогнозирования для оценки помех между станциями, находящимися на поверхности Земли, на частотах выше приблизительно 0,1 ГГц» [29]:

$$W_т = \frac{-H_{O_2} A_{O_2} - H_{H_2O} A_{H_2O} - H_{ГМ} A_{ГМ}}{\sin\beta}, \qquad (11)$$

где A_{O2}, A_{H2O}, $A_{ГМ}$ – коэффициенты погонного поглощения в кислороде, парах воды и гидрометеорах, соответственно (значения указанных величин в зависимости от частоты радиоизлучения, времени года, времени суток, урбанизированности местности; $H_{O2} \approx 5{,}3$ км, $H_{H2O} \approx 2{,}1$ км, $H_{ГМ} \approx 0{,}5\ldots2$ км – эквивалентная толщина слоев кислорода, водяных паров и гидрометеоров в радиолинии (см. рис. 4); β – угол места, определяющий направление радиолинии от ИРИ к приемнику сигнала (или наоборот).

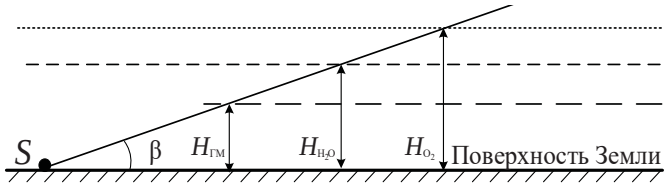


Рис. 4. Принцип расчета множителя ослабления на трассе

На практике расчет средней мощности сигнала на входе РПУ удобнее проводить, если все величины выражены в дБ. В этом случае формула для средней мощности сигнала (5) примет вид:

$$\bar{P}_c = P_1 + \eta_{ф1} + G_1 + W_0 + W_т + G_2 + \eta_{ф2} + \xi_c + \xi_п + \ldots \qquad (12)$$

Следует отметить, что справочные данные по значениям A_{O2}, A_{H2O}, $A_{ГМ}$ в единицах дБ/км приведены в [29].

Расчет средней мощности шума на входе РПУ.

На втором этапе методики выполняется расчет средней мощности шума $\bar{P}_ш$ на входе РПУ. Воздействующие на вход РПУ шумы в соответствии с природой их происхождения могут быть разделены на внутренние и внешние. К внутренним шумам следует отнести собственные тепловые шумы приемника, фидера и антенны. А к внешним шумам – атмосферные, электризационные, промышленные, станционные, космические, а также тепловое излучение атмосферы и Земли.

Количественно уровень (мощность) шумов оценивается величиной эффективной шумовой температуры $T_э$, в соответствии с формулой [24]:

$$P_ш = 10\lg(kBT_э). \qquad (13)$$

где $k = 1{,}38 \cdot 10^{-23}$ Вт/(Гц×K) – постоянная Больцмана; B – рабочая полоса частот приемника, Гц.

Эффективная шумовая температура $T_э$ определяется шумовой температурой внешних шумов T_A (шумовой температурой антенны), температурой фидера (обычно при расчетах принимаемой равной T_0) и температурой приемника $T_{пр}$

$$T_э = T_A \eta_2 \xi_c + T_0(1 - \eta_2 \xi_c) + T_{пр}. \qquad (14)$$

где $\eta_2 = \eta_{A2}\eta_{Ф2}$ – коэффициент полезного действия антенно-фидерного тракта приемника; $T_0 = 290\,K$ – стандартная абсолютная температура.

Шумовая температура приемника $T_{пр}$ соответствующая его пороговой чувствительности.

Шумовая температура антенны T_A определяется суммой шумовых температур различных источников помех

$$T_A = T_{A\,атм} + T_{A\,э} + T_{A\,пр} + T_{A\,ст} + T_{A\,к} + T_{A\,а} + T_{A\,з} + T_{A\,отр} \qquad (15)$$

где $T_{A\,атм}$ – шумовая температура атмосферных помех;

$T_{A\,э}$ – шумовая температура электризационных помех;

$T_{A\,пр}$ – шумовая температура промышленных помех;

$T_{A\,ст}$ – шумовая температура станционных помех;

$T_{A\,к}$ – шумовая температура космических помех;

$T_{A\,а}$ – шумовая температура теплового излучения атмосферы;

$T_{A\,з}$ – шумовая температура излучения Земли;

$T_{A\,отр}$ – шумовая температура отраженного от Земли излучения атмосферы ($T_{A\,з}+T_{A\,отр}\approx 290\ K$).

Источником атмосферных помех являются дальние грозовые разряды. Так, в течение суток протекает около 500 гроз (наибольшая грозовая деятельность свойственна тропическим районам). В среднем в мире происходит около 100 грозовых разрядов в секунду.

При этом велика роль и промышленных помех. Они почти не испытывают суточных и сезонных изменений, но их уровень меняется в зависимости от насыщенности района электроустановками и принятых мер по экранировке излучений.

Тепловое излучение Земли обусловлено собственным тепловым излучением, которое характерно для любого нагретого тела. Для ровной земной поверхности шумовая температура зависит от угла места β, вида поляризации и электрических параметров Земли.

Помимо создания собственного излучения поверхность Земли отражает тепловое излучение атмосферы.

Вычисление среднего ОСШ.

Расчет средних значений мощностей сигнала и помех позволяет получить среднее значение ОСШ $\overline{z}=\overline{P}_{с}/\overline{P}_{ш}$.

Вычисление СКО ОСШ.

Учитывая, что среднее значение ОСШ на входе РПУ является случайной величиной распределенной по нормальному закону, которая характеризуется двумя параметрами: математическим ожиданием (средним значением ОСШ $\overline{z}$) и среднеквадратическим отклонением (СКО), который учитывает разброс мгновенных значений z вокруг его среднего значения.

Величина СКО ОСШ определяется выражением

$$\sigma_z=\sqrt{\sigma_{с}^2+\sigma_{п}^2+\sigma_{пар}^2}\ . \qquad (16)$$

где $\sigma_{с}=2...4$ дБ – СКО мощности сигнала; $\sigma_{ш}=3...4$ дБ – СКО мощности шума; $\sigma_{пар}$ – СКО параметров радиоканала, определяемое выражением [18, 19]:

$$\sigma_{пар}=\sqrt{\sigma_{тр}^2+\sigma_{P_1}^2+\sigma_{АФТ_1}^2+\sigma_{АФТ_2}^2}\ . \qquad (17)$$

где $\sigma_{тр}=1...2$ дБ – СКО медленных изменений параметров радиотрассы: $\sigma_{P_1}=1...2$ дБ – СКО изменения мощности передатчика; $\sigma_{АФТ_1}=2...3$ дБ – СКО параметров передающего антенно-фидерного тракта; $\sigma_{АФТ_2}^2=2...3$ – СКО параметров приемного антенно-фидерного тракта.

Определение требуемого превышения мощности сигнала над мощностью шума.

Данный этап учитывает требуемое превышение, определяемое возможностью последующей обработки принятого сигнала. Учитывая, что в ГСС Starlink используются сигналы QAM-2, будем полагать, что на приеме требуемое превышение сигнала над помехой, с учетом помехоустойчивого кодирования, должно составляет $z_{тр}$ = 5...8 дБ.

Вычисление аргумента функции Лапласа.

Для оценки вероятности ЭМД предварительно вычисляют аргумент функции Лапласа $u=\dfrac{\overline{z}-z_{тр}}{\sigma_z}$.

Оценка вероятности ЭМД ИРИ.

Для получения численного значения вероятности ЭМД $P_{ЭМД}=\Phi(u)$ возможно использовать табулированное значение функции Лапласа, или воспользоваться графиком, представленным на рисунке 5, полученным в соответствии с формулой (4).

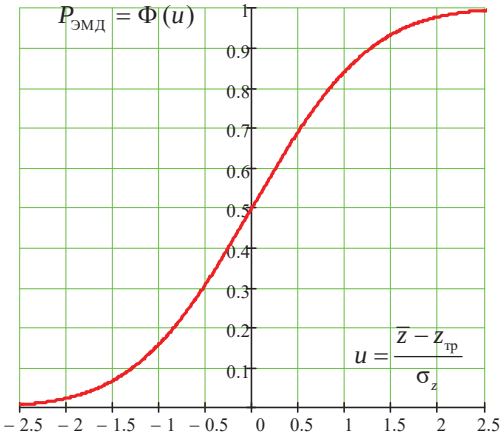


Рис. 5. Расчет вероятности электромагнитной доступности, в зависимости от аргумента *U*

Представленная методика оценки вероятности ЭМД адаптирована для ее применения к ИРИ ГСС Starlink.

Расчет оценка ЭМД ИРИ ГСС Starlink

На основе разработанной методики оценки вероятности ЭМД источников радиоизлучений ГСС Starlink возможно провести энергетические расчеты нисходящего канала (КА – АТ).

Рассматривается наиболее сложный случай, когда АТ Starlink находится в надире – в центре зоны покрытия КА и одновременно в центре луча, формируемого КА в направлении на АТ и направленного в надир. Тогда, исходя из особенностей работы ГССС Starlink, зона обслуживания луча диаграммы направленности КА, направленного на АТ будет иметь самый маленький диаметр – 33 км (см. рис. 1).

Во всех остальных случаях диаметр зоны обслуживания будет больше и на краю зоны покрытия КА будет достигать 222 км.

Очевидно, что с точки зрения радиомониторинга, попасть в зону обслуживания луча легче, когда он размещен на краю зоны покрытия КА, нежели чем он размещен в надире. Поэтому, если выполнить требования по ЭМД нисходящего канала (КА – АТ) для случая, когда АТ размещен в надире, то и для всех остальных конфигураций КА–АТ эти требования будут достаточными.

Рассматривается три варианта размещения наземного комплекса радиомониторинга (КРМ) относительно АТ (см. рис. 6):

вариант № 1 – K_1 КРМ на расстоянии $r = 10$ км от АТ Starlink *A*;

вариант № 2 – K_2 КРМ на расстоянии $r = 100$ км от АТ Starlink *A*;

вариант № 3 – K_3 КРМ на расстоянии $r = 900$ км от АТ Starlink *A*.

Исходя из принятых условий и ограничений в варианте № 1 КРМ размещен в зоне обслуживания луча диаграммы направленности КА, а в вариантах № 2 и 3 – за ее пределами.

Остальные исходные данные, а также заданные и рассчитанные энергетические характеристики нисходящего канала (КА – АТ) с позиции КРМ представлены в таблице 5.

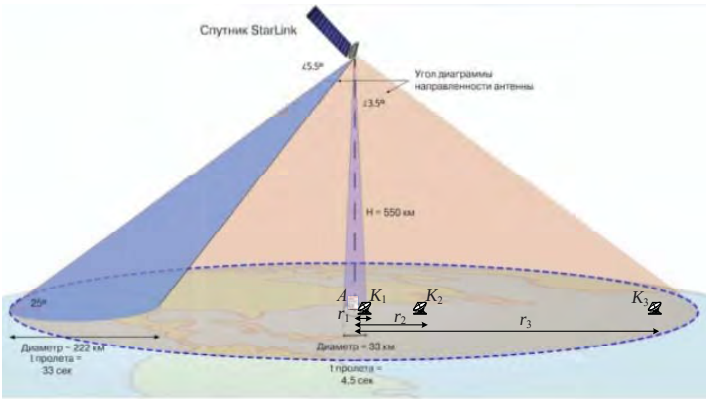


Рис. 6. Варианты размещения абонентских терминалов системы Starlink для луча, направленного в надир

Таблица 5

Исходные данные и рассчитанные значения

№ пп.	Характеристика	Значение характеристики		
		Вар. № 1	Вар. № 2	Вар. № 3
Исходные данные:				
1.	ЭИИМ КА Starlink, дБ	32		
2.	Рабочая частота нисходящий канала (КА – АТ) *f*, ГГц	12		
3.	Полоса радиосигнала, МГц	60		
4.	Удаление КРМ от АТ (расстояние по поверхности Земли) *d*, км	10	100	900
Этап 1. Расчет средней мощности сигнала на входе РПУ КРМ P_c				
5.	Наклонная дальность *r*, км	520	530	1070
6.	Угол места β, °	88	78	25
7.	Ослабление в свободном пространстве $W_0^2(r)$, дБ	–168,3	–168,5	–174,6
8.	Коэффициент согласования по сопротивлению ξ_c, дБ	–0,5		
9.	Коэффициент согласования по пространственной ориентации антенн $\xi_{пр}(\beta)$, дБ	–0,5	–30	–35,2
10.	Коэффициент согласования по поляризации $\xi_п$, дБ	–0,5		
11.	КПД фидера приемной антенны КРМ $\eta_{ф2}$, дБ	–0,5		
12.	Эквивалентная толщина слоя кислорода H_{O2}, км	5,3		
13.	Эквивалентная толщина слоя водяных паров H_{H2O}, км	2,1		
14.	Эквивалентная толщина слоя гидрометеоров $H_{ГМ}$, км	1,5		
15.	Коэффициент погонного поглощения в кислороде A_{O2}, дБ/км	0,006		
16.	Коэффициент погонного поглощения в парах воды A_{H2O}, дБ/км	0,006		
17.	Коэффициент погонного поглощения в гидрометеорах $A_{ГМ}$, дБ/км	0,1		
18.	Множитель ослабления сигнала на радиотрассе $W_т$, дБ	–0,3	–0,2	–0,4
19.	КИП раскрыва антенны	0,8		
20.	Диаметр антенн D_A, м	0,4…20*		
21.	Коэффициент усиления приемной антенны КРМ $G_2(D_A)$, дБ	33,1…67,1*		
22.	Средняя мощность сигнала на входе РПУ КРМ $P_c(D_A)$, дБ	–105,3… –71,3*	–134,9… –101,1*	–146,5… –112,5*
Этап 2. Расчет средней мощности шума на входе РПУ КРМ $P_ш$				
23.	Шумовая температура РПУ $T_{пр}$, *K*	80		
24.	Шумовая температура атмосферных помех $T_{A\,атм}$, *K*	45*	47**	70**
25.	Шумовая температура антенны T_A, *K*	81,8	83,5	103,1
26.	Эффективная шумовая температура $T_э$, *K*	171,5	173,2	191,8
27.	Средняя мощность шума на входе РПУ КРМ $P_ш$, дБ	–128,5	–128,4	–127,9
Этап 3. Вычисление среднего ОСШ				
28.	Среднее ОСШ, дБ	23,2… 57,1*	–6,6… 27,4*	–18,5… 15,5*
Этап 4. Вычисление СКО ОСШ σ_z				
29.	СКО медленных изменений параметров радиотрассы $\sigma_{тр}$, дБ	2		
30.	СКО изменения мощности передатчика σ_{P1}, дБ	2		
31.	СКО параметров передающего антенно-фидерного тракта $\sigma_{АФТ1}$, дБ	2,5		
32.	СКО параметров приемного антенно-фидерного тракта КРМ $\sigma_{АФТ2}$, дБ	3		
33.	СКО параметров радиоканала $\sigma_{пар}$, дБ	4,8		
34.	СКО мощности сигнала σ_c, дБ	3		
35.	СКО мощности шума $\sigma_ш$, дБ	3,5		
36.	СКО ОСШ σ_z, дБ	6,7		
Этап 5. Определение требуемого превышения мощности сигнала над мощностью шума $z_{тр}$				
37.	Требуемое превышение мощности сигнала над мощностью шума $z_{тр}$, дБ	10		
Этап 6. Вычисление аргумента функции Лапласа *u*				
38.	Аргумент функции Лапласа $u(D_A)$	1,97… 7,07*	–2,48… 2,61*	–4,28… 2,82*
Этап 7. Расчет вероятности ЭМД ИРИ $P_{ЭМД}$				
39.	Вероятность ЭМД ИРИ $P_{ЭМД}$	0,975… 0,9999*	0,007… 0,9955*	0,000… 0,7935*

*Расчет проводился для каждого значения диаметра антенны в интервале 0,4…20 м с шагом 0,1 м

**Значения шумовой температура теплового излучения атмосферы зависят от угла места, оценивались по графику [19]

Итоговые результаты расчета более наглядны в виде графиков зависимости вероятности ЭМД $P_{эмд}$ нисходящего канала (КА – АТ) с позиций КРМ от диаметра антенны D_A КРМ для вариантов № 1, 2, 3 (см. рис. 7).

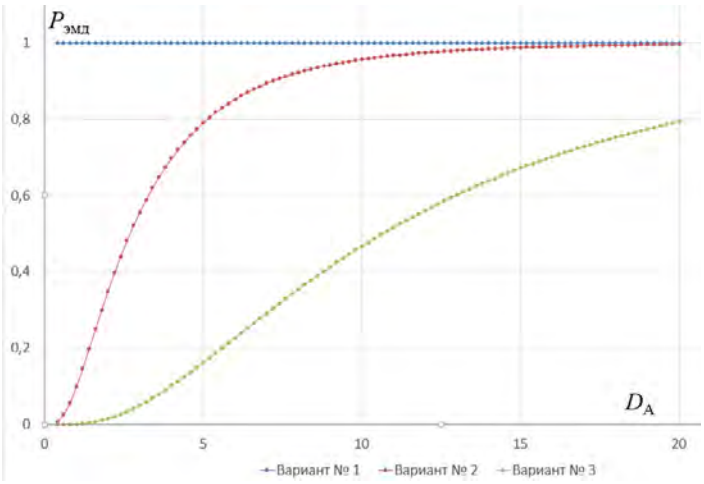


Рис. 7. Графики зависимости вероятности электромагнитной доступности $P_{эмд}$ КА Starlink для луча, направленного в надир от диаметра антенны D_A для вариантов № 1, 2, 3

Согласно полученным результатам, ЭМД ИРИ, находящихся на орбите, обеспечивается при удалении порядка 100 км от надира, где предположительно расположен АТ (около 530 км наклонной дальности). При этом следует понимать, что временной интервал, при котором достигается требуемая вероятность контакта в таких условиях, весьма непродолжителен. Применение низколетящих летательных аппаратах для решения такой задачи будет оправдано лишь при одновременном контроле АТ. Только при таких условиях обеспечивается контроль радиолинии.

Заключение

Разработка методики оценки вероятности ЭМД источников радиоизлучений ГСС Starlink и проведенные расчеты позволяют выработать некоторые рекомендации по повышению ЭМД ИРИ:

1. Применение антенн с хорошими направленными свойствами. Повышение коэффициента усиления (КУ) приемной антенны G_2 в рассматриваемой ситуации возможно лишь за счет увеличения диаметра отражателя, что позволяет существенно повысить мощность полезного сигнала на входе РПУ КРМ, а, следовательно, и результирующую вероятность ЭМД. Здесь необходимо обратить внимание на то, что при увеличении КУ сужается главный лепесток ДН антенной системы, что обуславливает ужесточения требований к точности ориентирования приемной антенны $\xi_п$ в направлении на КА.

2. ужение рабочей полосы частот РПУ КРМ. Анализ выражения, характеризующего расчет средней мощности шума показывает, что величина $P_ш$ прямо пропорциональна полосе тракта приема. Следовательно, сужение полосы приема ведет к снижению мощность шума, а значит повышению вероятности ЭМД. Однако сужать полосу приема B допустимо лишь в ограниченных пределах, ее нельзя делать меньше чем ширина спектра сигнала, в противном случае будут происходить его искажение. Еще большего снижения уровня шума возможно обеспечить за счет адаптивного согласования фильтра со значимой полосой спектра сигнала в ходе мониторинга. Тем самым обеспечивая совпадение передаточной функции фильтра с функцией огибающей спектра сигнала в каждый момент времени.

3. Использование малошумящих усилителей и приемников. Такая мера довольно частот используется в СВЧ-диапазоне, в котором собственные шумы приемного тракта становятся соизмеримы с внешними шумами. Учитывая, что основными источниками шумов в этом случае становятся электронные элементы усилительные каскады, то изначальный выбор элементной базы с низким уровнем коэффициента шума, является предпочтительным. К таким усилителям, например, относятся мазеры, охлаждаемые жидким азотом или гелием, параметрические усилители и др.

4. Использование КРМ, размещенных на летательных аппаратах. Такой подход открывает возможность пространственного размаха для выбора условий, при котором будет обеспечен одновременный контакт как космического аппарата, так и АТ. Если заранее известен район мониторинга, то возможно построить маршрут летательного аппарата таким образом, чтобы он смог залететь в зону обслуживания луча ДН КА. Тогда размер апертуры антенной системы летательного аппарата, обеспечивающего заданную вероятность ЭМД, может быть соизмерим с диаметром антенны АТ, т.е. достигать единицы десятков сантиметров.

Дальнейшие исследования авторы связывают с применением методов совместной частотно-временной обработки сигналов, предложенных в [30], при решении задач мониторинга.

Литература

1. *Пехтерев С.В., Макаренко С.И., Ковальский А.А.* Описательная модель системы спутниковой связи Starlink // Системы управления, связи и безопасности. 2022. № 4. С. 190-255. DOI 10.24412/2410-9916-2022-4-190-255.

2. *Паршин Г.К.* Проект Starlink: цели, реализация, перспективы // Наукоемкие технологии и интеллектуальные системы : Сборник статей по итогам Международной научно-практической конференции, Самара, 23 ноября 2018 года. Самара: Общество с ограниченной ответственностью "Агентство международных исследований", 2018. С. 50-52.

3. *Мырова Л.О., Ментус О.В., Давыдов А.Б.* и др. Низкоорбитальные спутниковые системы связи Starlink и OneWeb // Труды Научно-исследовательского института радио. 2021. № 2. С. 36-45. DOI 10.34832/NIIR.2021.5.2.005.

4. *Бестугин А.Р., Оводенко А.А., Крячко А.Ф., Киршина И.А.* Теория информационных управляющих комплексов на базе низкоорбитальных сетевых структур (монография). С-Пб: ГУАП (издана при финансовой поддержке РФФИ), 2015. 264 с.

5. *Козырев А.В., Дубинский А.А.* Спутниковая система Starlink // XLVII Гагаринские чтения 2021 : Сборник тезисов работ XLVII Международной молодёжной научной конференции, Москва, 20-23 апреля 2021 года. М.: Издательство "Перо", 2021. 608 с.

6. *Анпилогов В., Пехтерев С., Шишлов А.* Антенная решетка и абонентский терминал Starlink // Технологии и средства связи. 2020. № S1. С. 69-76.

7. *Игнатьев В.К., Перченко С.В., Станкевич Д.А.* Спиновый эффект Холла в поликристаллических образцах немагнитных металлов пятого и шестого периодов // Письма в Журнал технической физики. 2023. Т. 49, № 6. С. 25-27. DOI 10.21883/PJTF.2023.06.54812.19437.

8. Патент № 2740078 C1 Российская Федерация, МПК F03H 1/00, B64G 7/00. Ракетный лабораторный двигатель на эффекте Холла и стенд для его испытаний : № 2020124647 : заявл. 24.07.2020 : опубл. 11.01.2021 / А. С. Воронов, А. А. Троицкий, А. И. Стародубов ; заявитель Закрытое акционерное общество "СуперОкс" (ЗАО "СуперОкс").

9. *Висхаев М.Р., Албогачиева Л.А.* Глобальный спутниковый интернет Starlink // В сборнике: Приоритетные направления инновационной деятельности в промышленности. Сборник научных статей по итогам одиннадцатой международной научной конференции. 2020. С. 55-57.

10. *Пехтерев С.* Пропускная способность группировки StarLink в США и оценка ее потенциальной абонентской базы // Технологии и средства связи. 2021. № S1. С. 69-74.

11. *Анпилогов В., Пехтерев С., Шишлов А.* Анализ терминалов, планируемых для применения в системах Starlink и OneWeb // Технологии и средства связи. 2021. № S1. С. 30-36. EDN OYCOKN.

12. Starlink Satellite Constellation of SpaceX // eoPortal Directory [Электронный ресурс], 2021. URL: https://www.eoportal.org/satellite-missions/starlink (дата доступа 24.05.2023).

13. Starlink // Wikipedia [Электронный ресурс], 20.05.2023. URL: https://en.wikipedia.org/wiki/Starlink (дата доступа 24.05.2023)/

14. Starlink Slowed in Q2, Competitors Mounting Challenges // Ookla [Электронный ресурс], 20.05.2023. URL: https://www.ookla.com/articles/starlink-hughesnet-viasat-performance-q2-2022 (дата доступа 24.05.2023).

15. Receiving Starlink Beacons with an RTL-SDR and LNB [Электронный ресурс], 20.05.2023. URL: https://sgcderek.github.io/blog/starlink-beacons.html (дата доступа 24.05.2023).

16. *Пехтерев С.* Всё о проекте "Спутниковый интернет Starlink". Часть 1 // Хабр [Электронный ресурс], 20.05.2023. URL: https://habr.com/ru/post/526154/(дата доступа 24.05.2023).

17. *Дворников С.С., Устинов А.А., Дворников А.С.* и др. Модель энергетической доступности OFDM-каналов с сигналами BPSK // Вопросы радиоэлектроники. Серия: Техника телевидения. 2020. № 1. С. 98-103. EDN LZUHGH.

18. *Еремеев И.Ю., Демичев И В., Шайдулин З.Ф.* и др. Методика оценивания электромагнитной доступности в многопозиционных радиотехнических системах // Электромагнитные волны и электронные системы. 2022. Т. 27, № 2. С. 5-11. DOI 10.18127/j15604128-202202-01. EDN MRMXYZ.

19. *Липатников В.А., Царик О.В.* Методы радиоконтроля: теория и практика. СПб.: ГНИИ "Нацразвитие". 2018. 607 с.

20. *Баранов С.А., Шабров А.В.* Оценка зоны электромагнитной доступности радиоэлектронных средств при осуществлении радиоконтроля в декаметровом диапазоне радиочастот // Труды международного симпозиума "Надежность и качество". 2019. Т. 1. С. 251-253.

21. *Гордиенко Д.Ю., Дворников С.В.* Корреляционный прием частотно-манипулированных сигналов в режиме с псевдослучайной перестройкой рабочей частоты // T-Comm: Телекоммуникации и транспорт. 2022. Т. 16, № 3. С. 18-22. DOI 10.36724/2072-8735-2022-16-3-18-22.

22. *Ашихмин А.В., Виноградов А.Д., Рембовский А.М., Сладких В.А.* Способ однопозиционного местоопределения источников радиоизлучения с использованием бортового радиопеленгатора беспилотного летательного аппарата вертолетного типа // Системы управления, связи и безопасности. 2021. № 4. С. 40-57. DOI 10.24412/2410-9916-2021-4-40-57.

23. *Dvornikov S.S., Zheglov K.D., Dvornikov S.V.* SSB signals with controlled pilot level // T-Comm: Телекоммуникации и транспорт. 2023. Т. 17, № 3. С. 41-47. DOI 10.36724/2072-8735-2023-17-3-41-47.

24. *Попов О.В., Шепилов А.М., Борисов Г.Н.* и др. Методика расчета отношения мощностей собственных и внешних шумов на входе несогласованного приемника // Успехи современной радиоэлектроники. 2019. № 9. С. 52-57. DOI 10.18127/j20700784-201909-06.

25. *Дворников С.В., Дворников С.С., Жеглов К.Д.* Помехоустойчивость сигналов однополосной модуляции с управляемым уровнем несущего колебания // Информатика и автоматизация. 2023. Т. 22, № 2. С. 261-288. DOI 10.15622/ia.22.2.2.

26. *Вознюк В.В., Куценко Е.В., Ворона С.Г.* Исследование потенциальной помехоустойчивости оптимального приемника с возможностью адаптации к виду и параметрам помех в условиях воздействия множества узкополосных шумовых помех // Нелинейный мир. 2020. Т. 18, № 4. С. 41-57. DOI 10.18127/j20700970-202004-05.

27. *Антохин Е.А., Власенко В.И., Бестугин А.Р.* и др. Синтез широкодиапазонных однопроводных антенн бегущей волны // Радиотехника. 2022. Т. 86, № 7. С. 142-151. DOI 10.18127/j00338486-202207-20.

28. *Манаенко С.С., Дворников С.В., Пшеничников А.В.* Теоретические аспекты формирования сигнальных конструкций сложной структуры // Информатика и автоматизация. 2022. Т. 21, № 1. С. 68-94. DOI 10.15622/ia.2022.21.3.

29. Рекомендация МСЭ-R P.452-15 (09/2013) Процедура прогнозирования для оценки помех между станциями, находящимися на поверхности Земли, на частотах выше приблизительно 0,1 ГГц. Серия Р. Распространение радиоволн. [Электронный ресурс], 20.05.2023. http://www.itu.int/ITU-R/go/patents/en, (дата доступа 24.05.2023).

30. *Дворников С.В., Марков Е.В., Маноши Э.А.* Повышение помехозащищенности передач декаметровых радиоканалов в условиях непреднамеренных помех // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15, № 6. С. 4-9. DOI 10.36724/2072-8735-2021-15-6-4-9.

EVALUATION OF ELECTROMAGNETIC AVAILABILITY OF RADIO EMISSION SOURCES STARLINK GSS

VLADIMIR V. SEVIDOV,
Saint-Petersburg, Russia

SERGEY S. DVORNIKOV,
Saint-Petersburg, Russia

ALEXANDER R. BESTUGIN,
Saint-Petersburg, Russia

SERGEY V. DVORNIKOV,
Saint-Petersburg, Russia

IRINA A. KIRSHINA,
Saint-Petersburg, Russia

KEYWORDS: *radio monitoring of Starlink system subscribers, electromagnetic availability of radio emission sources, method of assessing the electromagnetic availability of subscribers in Ku/K-frequency bands, noise immunity of Starlink system signal reception.*

ABSTRACT

Introduction: The full-scale deployment of the Starlink global satellite system has opened up wide opportunities for high-speed wireless access to Internet information resources throughout the Earth. In such conditions, the problem of user control over the provided traffic arises. This is due to the fact that the organizational and technical parameters of the Starlink system were initially chosen with the condition of minimizing mutual interference to subscribers located in limited areas. On the one hand, such a system opens up the prospect of access to the global network for subscribers located in remote, industrially unequipped areas of the globe, and on the other hand, the possibility of organizing communications for illegitimate users and the right of criminal organizations. Therefore, it is necessary to understand how accessible the communication channels organized by the Starlink system are for monitoring. Taking into account these circumstances, the purpose of the study, the materials of which are presented in this article, was to assess the probability of electromagnetic availability of subscribers of the Starlink global satellite system. Based on the analysis of the operational and technical characteristics of the Starlink system in available sources, the features of its construction were considered and the initial data were determined for the development of a method for the electromagnetic availability of radio emission sources operating in the Ku / K-frequency bands. **The methodology** is based on calculations of the provided excess of signal power over noise power at the input of the radio receiver, which ensures the required quality and reliability of reception. The stages of the methodology are substantiated, the analytical apparatus is modified, taking into account the conditions of monitoring. **The results** of calculating the electromagnetic availability of radio emission sources of the Starlink global satellite system for the most difficult case, when the subscriber terminal is located at the nadir of the spacecraft coverage area and at the same time in the center of the beam formed on the subscriber, are presented. Probabilistic estimates are obtained and proposals and recommendations are formulated to improve the reliability and quality of the ongoing monitoring. Conclusions and proposals for the practical implementation of the results obtained are formulated.

REFERENCES

1. S.V. Pekhterev, S.I. Makarenko, A.A. Kovalsky, “Descriptive model of the Starlink satellite communication system,” *Control, communication and security systems*. 2022. No. 4, pp. 190-255. DOI 10.24412/2410-9916-2022-4-190-255.

2. G.K. Parshin, “Starlink project: goals, implementation, prospects,” *High technologies and intelligent systems: Collection of articles based on the results of the International Scientific and Practical Conference*, Samara, November 23, 2018. Samara: Limited Liability Company “Agency for International Research”, 2018, pp. 50-52.

3. L.O. Myrova, O.V. Mentus, A.B. Davydov etc., “Low-orbit satellite communication systems Starlink and OneWeb,” *Proceedings of the Radio Research Institute*. 2021. No. 2, pp. 36-45. DOI 10.34832/NIIR.2021.5.2.005.

4. A.R. Bestugin, A.A. Ovodenko, A.F. Kryachko, I.A. Kirshina, “Theory of information control complexes based on low-orbit network structures (monograph),” St. Petersburg: GUAP, 2015. 264 p.

5. A.V. Kozyrev, A.A. Dubinsky, “Starlink satellite system,” *XLVII Gagarin Readings 2021: Collection of abstracts of the XLVII International Youth Scientific Conference*, Moscow, April 20-23, 2021. Moscow: Pero Publishing House, 2021. 608 p.

6. V. Anpilogov, S. Pekhterev, A. Shishlov, “Antenna array and Starlink subscriber terminal,” *Technologies and communications*. 2020. No. S1, pp. 69-76.

7. V.K. Ignatiev, S.V. Perchenko, D.A. Stankevich, “Spin Hall effect in polycrystalline samples of non-magnetic metals of the fifth and sixth periods,” *Letters to the Journal of Technical Physics*. 2023. Vol. 49, No. 6, pp. 25-27. DOI 10.21883/PJTF.2023.06.54812.19437.

8. Patent No. 2740078 C1 Russian Federation, IPC F03H 1/00, B64G 7/00. Laboratory rocket engine based on the Hall effect and a stand for its testing: No. 2020124647: application. 07/24/2020: publ. 01/11/2021 / A. S. Voronov, A. A. Troitsky, A. I. Starodubov; applicant Closed Joint Stock Company "SuperOx" (ZAO "SuperOx").

9. M.R. Vikhaev, L.A. Albogachieva, “Global satellite Internet Starlink,” *Priority directions of innovation activity in industry. Collection of scientific articles based on the results of the eleventh international scientific conference*. 2020, pp. 55-57.

10. S. Pekhterev, “Bandwidth capacity of the StarLink group in the USA and assessment of its potential subscriber base,” *Technologies and communications*. 2021. No. S1, pp. 69-74.

11. V. Anpilogov, S. Pekhterev, A. Shishlov, “Analysis of terminals planned for use in the Starlink and OneWeb systems,” *Technologies and communications*. 2021. No. S1, pp. 30-36.

12. Starlink Satellite Constellation of SpaceX // eoPortal Directory [Electronic resource], 2021. URL: https://www.eoportal.org/satellite-missions/starlink (access date 05/24/2023).

13. Starlink // Wikipedia [Electronic resource], 05/20/2023. URL: https://en.wikipedia.org/wiki/Starlink (accessed 05/24/2023)/

14. Starlink Slowed in Q2, Competitors Mounting Challenges // Ookla [Electronic resource], 05.20.2023. URL: https://www.ookla.com/articles/starlink-hughesnet-viasat-performance-q2-2022 (accessed 05/24/2023).

15. Receiving Starlink Beacons with an RTL-SDR and LNB [Electronic resource], 05.20.2023. URL: https://sgcderek.github.io/blog/starlink-beacons.html (accessed 05/24/2023).

16. S. Pekhterev, "All about the Starlink Satellite Internet project. Part," Habr [Electronic resource], 05.20.2023. URL: https://habr.com/ru/post/526154/(access date 05/24/2023).

17. S.S. Dvornikov, A.A. Ustinov, A.S. Dvornikov et al., "Model of energy availability of OFDM channels with BPSK signals," *Problems of radio electronics. Series: Television technology*. 2020. No. 1, pp. 98-103.

18. I.Yu. Eremeev, I.V. Demichev, Z.F. Shaidulin et al., "Methodology for assessing electromagnetic accessibility in multi-position radio systems," *Electromagnetic waves and electronic systems*. 2022. Vol. 27. No. 2, pp. 5-11. DOI 10.18127/j15604128-202202-01.

19. V.A. Lipatnikov, O.V. Tsarik, "Radio monitoring methods: theory and practice," SPb.: State Research Institute "National Development". 2018. 607 p.

20. S.A. Baranov, A.V. Shabrov, "Assessment of the electromagnetic accessibility zone of radio-electronic equipment when carrying out radio monitoring in the decameter radio frequency range," *Proceedings of the international symposium "Reliability and Quality"*. 2019. Vol. 1, pp. 251-253.

21. D.Yu. Gordienko, S.V. Dvornikov, "Correlation reception of frequency-keyed signals in a mode with pseudo-random tuning of the operating frequency," *T-Comm*. 2022. Vol. 16, No. 3, pp. 18-22. DOI 10.36724/2072-8735-2022-16-3-18-22.

22. A.V. Ashikhmin, A.D. Vinogradov, A.M. Rembovsky, V.A. Sladkikhv, "A method for single-position location of radio emission sources using an on-board radio direction finder of a helicopter-type unmanned aerial vehicle," *Control, communication and security systems*. 2021. No. 4, pp. 40-57. DOI 10.24412/2410-9916-2021-4-40-57.

23. S.S. Dvornikov, K.D. Zheglov, S.V. Dvornikov, "SSB signals with controlled pilot level," *T-Comm*. 2023. Vol. 17, No. 3, pp. 41-47. DOI 10.36724/2072-8735-2023-17-3-41-47.

24. O.V. Popov, A.M. Shepilov, G.N. Borisov etc., "Methodology for calculating the ratio of the powers of own and external noise at the input of an unmatched receiver," *Advances in modern radio electronics*. 2019. No. 9, pp. 52-57. DOI 10.18127/j20700784-201909-06.

25. S.V. Dvornikov, S.S. Dvornikov, K.D. Zheglov, "Noise immunity of single-sideband modulation signals with a controlled level of carrier oscillation," *Informatics and Automation*. 2023. Vol. 22, No. 2, pp. 261-288. DOI 10.15622/ia.22.2.2.

26. V.V. Voznyuk, E.V. Kutsenko, S.G. Vorona, "Study of the potential noise immunity of an optimal receiver with the ability to adapt to the type and parameters of interference under the influence of many narrow-band noise interferencem," *Nonlinear World*. 2020. Vol. 18, No. 4, pp. 41-57. DOI 10.18127/j20700970-202004-05.

27. E.A. Antokhin, V.I. Vlasenko, A.R. Bestugin etc., "Synthesis of wide-range single-wire traveling wave antennas," *Radio engineering*. 2022. Vol. 86, No. 7, pp. 142-151. DOI 10.18127/j00338486-202207-20.

28. S.S. Manaenko, S.V. Dvornikov, A.V. Pshenichnikov, "Theoretical aspects of the formation of signal structures of complex structure," *Informatics and Automation*. 2022. Vol. 21, No. 1, pp. 68-94. DOI 10.15622/ia.2022.21.3.

29. Recommendation ITU-R P.452-15 (09/2013) Prediction procedure for assessing interference between stations on the surface of the Earth at frequencies above approximately 0.1 GHz. Series P. Radio wave propagation. [Electronic resource], 05/20/2023. http://www.itu.int/ITU-R/go/patents/en, (accessed 05/24/2023).

30. S.V. Dvornikov, E.V. Markov, E.A. Manoshi, "Increasing the noise immunity of transmissions of decameter radio channels in conditions of unintentional interference," *T-Comm*, 2021. Vol. 15, No. 6, pp. 4-9. DOI 10.36724/2072-8735-2021-15-6-4-9.

INFORMATION ABOUT AUTHORS:

Vladimir V. Sevidov, Ph.D. tech. Sciences, Associate Professor, Doctoral student, Military Academy of Communications. CM. Budyonny. Research interests: theory of signal transmission, spectral efficiency of signals, noise immunity of control and communication channels of radio engineering systems, coordination of radio emission sources, St. Petersburg, Russia, v-v-sevidov@mail.ru, orcid.org/0009-0009-2413-1615

Sergey S. Dvornikov, Ph.D. tech. Sciences, Associate Professor, Department of Design and Technologies of Electronic and Laser Means, Federal State Autonomous Educational Institution of Higher Education "St. Petersburg State University of Aerospace Instrumentation"; Research Fellow, Research Department, Military Academy of Communications. CM. Budyonny. Research interests: theory of signal transmission, spectral efficiency of signals, noise immunity of control and communication channels of radio engineering systems, St. Petersburg, Russia, dvornik.92@mail.ru, orcid.org/0000-0001-7426-6475

Alexander R. Bestugin, Doctor of Science (Engineering), Professor, Director of the Institute of Radio Engineering and Infocommunication Technologies, Federal State Autonomous Educational Institution of Higher Education "Saint-Petersburg State University of Aerospace Instrumentation". Area of scientific interests: construction of noise-protected radio-communication systems, formation and processing of signals of complex structures, St. Petersburg, Russia, fresguap@mail.ru, orcid.org/0000-0003-3847-2516

Sergey V. Dvornikov, Doctor of Engineering. Sciences, Professor, Department of Radio Engineering and Optoelectronic Complexes, Federal State Autonomous Educational Institution of Higher Education "St. Petersburg State University of Aerospace Instrumentation"; Professor of the Department of Radio Communications, Military Academy of Communications. CM. Budyonny. Area of scientific interests: construction of noise-protected radio communication systems, formation and processing of signals of complex structures, St. Petersburg, Russia, practicdsv@yandex.ru, orcid.org/0000-0001-7426-6475

Irina A. Kirshina, Candidate of Science (Economics), Associate Professor, Department of Design and Technology of Electronic and Laser Devices, Federal State Autonomous Educational Institution of Higher Education "Saint-Petersburg State University of Aerospace Instrumentation. Research interests: theory of signal transmission, spectral efficiency of signals, interference immunity of control and communication channels of radio engineering systems. St. Petersburg, Russia, ikirshina@mail.ru, orcid.org/0000-0002-31102363

For citation: Sevidov V. V., Dvornikov S. S., Bestugin A. R., Kirshina I. A., Dvornikov S. V. Evaluation of the electromagnetic accessibility of radio emission sources Starlink GSS. *H&ES Reserch*. 2023. Vol. 15. No. 4. P. 26-37. doi: 10.36724/2409-5419-2023-15-4-26-37 (In Rus)

doi: 10.36724/2409-5419-2023-15-4-38-46

АНАЛИЗ БЕЗОПАСНОСТИ СИСТЕМ NOMA

АЛЕКСЕЕВ
Сергей Сергеевич[1]

КОСИЧКИНА
Татьяна Павловна[2]

ПАНКРАТОВ
Денис Юрьевич[3]

ШАМСУТДИНОВ
Илья Артурович[4]

АННОТАЦИЯ

Введение: Существующие схемы безопасности физического уровня (PLS, Physical Layer Security) для NOMA или базируются на подходах на основе криптографии, или ограничены подходами, которые требуют высокой обработки с вычислительной сложностью и совместного использования ключей. Обычная система NOMA страдает от рисков и недостатков безопасности, таких как склонность к внешнему или внутреннему прослушиванию. **Методы:** В результате рассылки сообщений NOMA нескольким пользователям в одно и то же время на одних и тех же ресурсах существует риск того, что неавторизованный пользователь может прослушивать или получать доступ к информации нескольких пользователей при условии успешного перехвата передачи NOMA. Система NOMA подвержена внутреннему перехвату при защите конфиденциальных данных в случаях присутствия ненадежных пользователей. Поэтому традиционные методы обмена данными не могут обеспечить необходимый уровень секретности систем NOMA. **Результаты исследования:** В работе рассматривается технология неортогонального доступа (NOMA) и оценка безопасности различных систем беспроводной связи с использованием NOMA, таких как когнитивные радиосети, систем с поддержкой ретрансляции для режима IoT, а также режима D2D. Приводятся результаты анализа характеристик безопасности и работы системы NOMA в разных режимах.

Сведения об авторах:

[1] МТУСИ, студент ЗРС1701, Москва, Россия, saleks00@yandex.ru

[2] МТУСИ, к.т.н. доц. каф. СиСРТ, Москва, Россия, t.p.kosichkina@mtuci.ru

[3] МТУСИ, к.т.н. доц. каф. СиСРТ, Москва, Россияа, dpankr@mail.ru

[4] МТУСИ, студент МИТ2201, Москва, Россия, malon228@mail.ru

КЛЮЧЕВЫЕ СЛОВА: *NOMA, безопасность, PLS, MIMO-NOMA, IoT, D2D.*

Для цитирования: *Алексеев С.С., Косичкина Т.П., Панкратов Д.Ю., Шамсутдинов И.А.* Анализ безопасности систем NOMA // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 38-46. doi: 10.36724/2409-5419-2023-15-4-38-46

Введение

Из-за широковещательного характера беспроводной связи безопасность сети всегда страдает от рисков перехвата, что может повлиять на конфиденциальность и безопасность сигналов. Существующие схемы безопасности физического уровня (PLS, Physical Layer Security) для NOMA или базируются на подходах на основе криптографии, или ограничены подходами, которые требуют высокой обработки с вычислительной сложностью и совместного использования ключей, что делает их неподходящими [1]. Обычная система NOMA страдает от рисков и недостатков безопасности, таких как склонность к внешнему или внутреннему прослушиванию. В результате рассылки сообщений NOMA нескольким пользователям в одно и то же время на одних и тех же ресурсах существует риск того, что внешний перехватчик (т.е. неавторизованный пользователь) может прослушивать или получать доступ к информации нескольких пользователей при условии успешного перехвата передачи NOMA. Кроме того, система NOMA подвержена внутреннему перехвату при защите конфиденциальных данных в случаях присутствия ненадежных пользователей. Поэтому традиционные методы обмена данными не могут обеспечить необходимый уровень секретности систем NOMA.

Технология NOMA

Сеть беспроводной связи пятого поколения (5G) представляет собой эволюцию предыдущих сетей четвертого поколения (4G) с несколькими новыми расширенными услугами [2], повышенной надежностью за пределами Интернета для критически важных коммуникаций и Интернетом вещей (IoT). Неортогональный множественный доступ (NOMA, Non-Orthogonal Multiple Access) был задуман как прорывная технология в 5G из-за его превосходной спектральной эффективности. В отличие от обычного ортогонального множественного доступа, NOMA обеспечивает более высокую пропускную способность и лучшую энергоэффективность, а также поддерживает массовое подключение, позволяя пользователям использовать одни и те же временные, частотные и кодовые ресурсы для передачи информации.

Скоординированная многоточечная передача (CoMP, Coordinated Multi-Point) является одним из многообещающих улучшений благодаря ее способности улучшать охват услуг с высокой скоростью передачи данных, увеличивать пропускную способность и контролировать уровень помех. Недавние исследования показали, что путем внедрения NOMA в сетях CoMP можно не только еще больше повысить эффективность использования ресурсов всей сети, но и уменьшить сложность реализации, используя правильную стратегию пользовательского планирования.

Несмотря на огромный объем исследований в области NOMA, очень немногие существующие исследования сосредоточены на вопросах безопасности. В частности, исследованы такие вопросы, как:

– оптимизация порядка декодирования сообщений для NOMA из соображений секретности;

– безопасность физического уровня NOMA в крупномасштабных сетях;

– метод формирования лучей диаграммы направленности с использованием искусственного шума для защиты личных данных двух пользователей в сети NOMA (эта модель разработана для систем MISO-NOMA, в которых перехватчик получает ухудшенные версии сигналов абонентов);

– исследование безопасности на физическом уровне, где легитимные и перехватывающие узлы моделируются с использованием стохастической геометрии.

Проблемы безопасности NOMA требуют дальнейших исследований и использования технологий, которые позволят обеспечить безопасность пользователей, соблюдая при этом требования к пропускной способности и эффективности сетей 5G. В связи с этим, в данной работе предлагается произвести анализ технологии NOMA и ее безопасности.

Анализ безопасности систем с NOMA

Секретность на физическом уровне в когнитивных радиосетях с поддержкой NOMA.

Базовая когнитивная радиосеть (Cognitive Radio Network, CRN) с неортогональным множественным доступом (NOMA) является многообещающей схемой множественного доступа для решения проблемы нехватки спектра. Эта новая базовая сеть CRN с поддержкой NOMA также может повысить секретность передачи за счет использования преднамеренного введения помех. Рассматривая существующего в сети перехватчика данных, было получено [3] выражение в закрытой форме для коэффициента суммы секретности (Secrecy Sum Rate, SSR) всех вторичных пользователей (Secondary User, SU). Затем формулируется задача оптимизации SSR как для основного пользователя (Primary User, PU), так и для SU, и разрабатывается алгоритм (Simulated Annealing Algorithm), чтобы найти оптимальное распределение мощности. Результаты моделирования демонстрируют, что показатели секретности всех SU с поддержкой NOMA выше, чем у SU с множественным доступом с частотным разделением (FDMA). Другими словами, использование технологии NOMA может повысить секретность сети CRN [3].

С распространением интеллектуальных мобильных устройств и соответствующих приложений в Интернете вещей (IoT), радио спектр быстро стал дефицитным и дорогим ресурсом [7]. Эффективные стратегии распределения спектра важны для реагирования на ограниченные ресурсы спектра в широко распределенных беспроводных сетях [8]. Есть два вида новых технологий исследования для удовлетворения потребностей в ресурсах спектра. Одним из них является технология когнитивного радио (CRN), которая может динамически выбирать каналы без межпользовательских помех.

Другое привлекательное решение – технология NOMA, которая позволяет нескольким пользователям передавать сигналы, используя один и тот же временной интервал. Вышеуказанные технологии для повышения спектральной эффективности основаны на совместном использовании спектра, что может привести к проблемам безопасности во время связи [9]. Например, хотя вторичный пользователь (SU) в когнитивных радиосетях (CRN) может получить доступ к спектру основного пользователя (PU), он также может прослушивать легальные сигналы из-за отсутствия соответствующей

политики управления использования спектра. Соответственно, были исследованы несколько моделей для обеспечения безопасности и качества передачи в когнитивных радиосетях (CRN) [4]–[6].

Ли и соавторы изучили стратегии повышения эффективности использования спектра для кооперативных сетей CRN с несколькими входами и несколькими выходами (MIMO) и проанализировали потенциал секретности на основе совместного глушения [4]. Авторы [5] изучали кооперативную безопасную связь для сети когнитивного радио с четырьмя узлами и получили достижимую скорость SU. Более того, если предположить, что SU не мешает передаче PU, [6] то в этом случае работает безопасный механизм совместной передачи для обоих PU и SU. Тем не менее, ясно, что скорость передачи данных SU с плохим состоянием канала может быть небольшой. В результате спектральная эффективность сети CRN с обычными методами множественного доступа невелика, поскольку ресурсы SU не могут быть доступны другим. К счастью, технология NOMA может гарантировать качество передачи пользователей с плохим каналом [10], [11] и разрешить большему количеству пользователей доступ к одному блоку ресурсов спектра одновременно [12].

В [13] Дин и соавторы, проанализировав полученную информацию о качестве двух систем NOMA (т.е. NOMA с фиксированным распределением мощности и NOMA в стиле CR), доказали, что NOMA с использованием CR, может гарантировать хорошее качество связи для пользователя с плохим состоянием канала. По сравнению с NOMA с фиксированным распределением мощности, схема CR NOMA может добиться большей эффективности при распределении спектра. Кроме того, авторы [14] разработали распределение мощности для двух SU в сети CRN с поддержкой NOMA для улучшения спектральной эффективности. Таким образом, это естественно применять NOMA в сетях CRN для значительного улучшения эффективности использования спектра.

В недавних исследованиях больше внимания уделялось развитию эффективных стратегий распределения спектра для сетей CRN с поддержкой NOMA. Принимая во внимание проблемы прослушивания телефонных разговоров в беспроводной связи, исследователи предлагают использовать совместное глушение для обеспечения безопасной передачи по физическому уровню в сети CRN с поддержкой NOMA [15]. Авторы в [16] проанализировали характеристики секретности сетей на основе NOMA новыми точными и асимптотическими выражениями для вероятности нарушения безопасности связи одного пользователя. Для случая системы NOMA с одной антенной (SISO), Чжан и др. представили алгоритм оптимального распределения мощности для максимизации коэффициента секретности (SSR, Secrecy Sum Rate) в [17]. В [18] авторы вывели новые выражения вероятности сбоя в замкнутой форме и проанализировали производительность в крупномасштабной сети на основе CRN с поддержкой NOMA и случайным образом расположенными пользователями. Хотя упомянутые выше источники в некоторой степени проливают свет на исследования в области безопасности NOMA, универсальное решение защищенной передачи для NOMA до сих пор не ясно.

Безопасная передача с помощью технологии формирования лучей для Интернета вещей с поддержкой ретрансляции.

Еще одной интересной является область защищенной передачи данных по нисходящей линии связи, которая открыта для многократного прослушивания, для приложений Интернета вещей (IoT) [32]. Предполагается наихудший сценарий в том смысле, что для повышения способности к перехвату все подслушивающие устройства расположены близко друг к другу и вступают в сговор с целью формирования совместного луча приема.

Для такой системы предложена новая схема защищенной передачи с совместным неортогональным множественным доступом (NOMA), для которой устройство Интернета вещей с более сильным каналом действует как ретранслятор энергии, чтобы помочь второму устройству Интернета вещей, работающему в условиях более слабого канала, и производительность проанализирована и оценена [32].

Сформулирована и решена задача максимизации коэффициента секретности (SSR) при трех ограничениях: 1) мощность передачи; 2) последовательное подавление помех; 3) качество обслуживания. Рассматривая сценарии как пассивного, так и активного подслушивания, предлагаются две схемы оптимизации для улучшения общего коэффициента SSR системы. С одной стороны, для сценария пассивного подслушивания предлагается защищенная схема формирования лучей с помощью искусственного шума. С другой стороны, для сценария активных подслушивающих устройств с несколькими антеннами рассматривается схема формирования лучей на основе ортогональной проекции.

Поскольку для передачи с одной антенной схема, основанная на ортогональной проекции, может быть неприменима, предлагается простая схема управления мощностью. Различные результаты оценки производительности, полученные с помощью компьютерного моделирования, подтвердили, что предлагаемые схемы превосходят другие эталонные схемы с точки зрения характеристик SSR [32].

Интернет вещей (IoT) быстро развивается как сложная платформа, соединяющая очень большое количество коммуникационных устройств, например, датчиков, контроллеров и др. [19]. Однако достижение требуемой повсеместной возможности подключения, необходимой для таких систем связи на основе IoT, становится жизненно важной и сложной задачей, главным образом из-за ограниченной пропускной способности. В этом контексте неортогональный множественный доступ (NOMA) предлагается как многообещающий метод поддержки концепции всепроникающей связи, позволяющий значительно повысить спектральную эффективность систем IoT [20-22], [12], [23].

Ключевой особенностью NOMA является реализация множественного доступа в области мощности, в то время как временные/частотные/кодовые ресурсы могут одновременно использоваться всеми пользователями. Более того, по сравнению с ортогональным доступом (OMA), доступ NOMA может обеспечить лучший баланс между суммарной скоростью и справедливой передачей данных пользователей [24]. В то же время, поскольку беспроводная природа распространения ра-

диосигналов делает связь IoT восприимчивой к атакам с целью прослушивания, при разработке таких систем необходимо очень тщательно учитывать аспекты их безопасности.

В прошлом для предотвращения перехвата секретных сообщений злоумышленниками применялись обычные методы шифрования [25], [26]. Однако таким методам шифрования свойственны трудности и уязвимости, связанные с управлением секретными ключами [19]. К счастью, безопасность физического уровня (PLS) продемонстрировала большой потенциал для более точного различения сигналов, принадлежащих законному получателю и перехватчику [27-29]. В отличие от методов, основанных на шифровании, методы PLS используют физические характеристики беспроводной среды для обеспечения информационной безопасности, независимой от вычислительных возможностей перехватчика [30], [31].

Стремясь еще больше повысить секретность систем 5G, Тиан и др. [20] объединили NOMA с методами использования нескольких антенн для оптимизации коэффициента секретности (SSR) беспроводной системы 5G для приложений, в которых доступна информация о состоянии канала (CSI) перехватчиков. В другом подходе Ху и др. [19] объединили метод передачи с несколькими антеннами с поддержкой искусственного шума (Artificial Noise, AN) с совместным глушением (Cooperative Jamming, CJ), чтобы уменьшить влияние пассивных перехватчиков на нисходящей линии IoT.

Поскольку многие передатчики устройств IoT имеют ограниченную мощность, выделение некоторой мощности для сигналов AN или помех может быть нецелесообразным, поскольку это ограничит зону покрытия безопасных передач. Чтобы заполнить этот пробел, в этой статье [32] предлагается новая схема безопасной передачи с помощью ретрансляции, предназначенная для систем связи на основе IoT с ограниченным энергопотреблением, где контроллер передает секретные сообщения двум классам устройств, работающих в присутствии нескольких подслушивающих устройств. Эти IoT-устройства предъявляют разнообразные требования к качеству обслуживания (QoS), а условия каналов между ними сильно различаются, т.е. условия канала для D1 намного лучше, чем условия канала для D2.

Чтобы гарантировать конфиденциальность передачи данных и требования QoS для D2, абонент D1 действует как ретранслятор энергии, чтобы помочь D2. Предполагая, что местонахождение всех подслушивающих устройств близко к контроллеру, этот наихудший сценарий позиционирования повысит вероятность перехвата информационного сигнала. Для этой системы IoT с ограничением энергии предлагается новая схема безопасной совместной передачи. Предполагая, что имеется информация CSI для подслушивающих устройств, рассматриваются два сценария работы, а именно: пассивные подслушивающие и активные подслушивающие устройства, для которых предлагается два способа оптимизации для максимизации коэффициента SSR рассматриваемой системы IoT при ограниченной мощности передачи, заданном QoS и ограничениях последовательного подавления помех (SIC).

На рисунке 1 жирными буквами обозначены векторы, характеризующие передачу информации от контроллера *S* с антенной системой двум пользовательским устройствами *Di*, $i \in \{1, 2\}$, каждое из которых оснащено несколькими антеннами. Контроллер S передает конфиденциальную информацию на устройства *Di* в присутствии подслушивателей (Eves). Также предполагается, что устройство D1, работающее в лучших условиях канала, используется для приема довольно коротких пакетов данных устройства D2, которое работает в условиях слабого канала, и представляет собой устройство, которое выполняет некоторые фоновые задачи, такие как загрузка мультимедийных файлов [32].

Как показано на рис. 1, подслушиватели G1e расположены близко друг к другу, и все они расположены близко к контроллеру S. Очевидно, что такая конфигурация приведет к более высокой вероятности перехвата передаваемого информационного сигнала. Для обеспечения высокой степени секретности для D2, D1 действует как ретранслятор энергии для пересылки несущего информацию сигнала на D2. Более конкретно, полученный сигнал на D1 разделяется на две части: одна для сбора энергии, а другая для декодирования информации, так что в совместной безопасной передаче NOMA будут задействованы два этапа. На стороне D2 принятые сигналы объединяются по схеме MRC (Maximum Ratio Combining). Предполагается, что во всех каналах с замираниями, показанных на рисунке 1, действуют независимые квазистатические замирания, которые остаются постоянными на одном временном интервале, но изменяются независимо от одного временного интервала к другому [32].

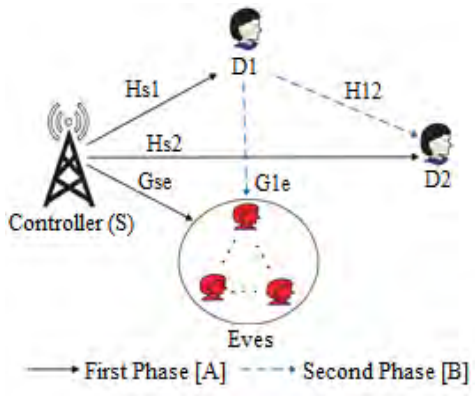


Рис. 1. Рассматриваемая модель безопасной совместной системы связи на основе IoT

Анализ вопросов повышения безопасности для кооперативных сетей D2D на базе технологии NOMA.

В статье исследователей школы информационных и коммуникационных технологий Сианьского университета Цзяотун (г. Сиань, Китай) на тему «Повышение безопасности и обеспечение качества обслуживания для кооперативных сетей D2D на базе NOMA» [33] исследуются характеристики гетерогенного использования энергии в сетях сотовой связи с поддержкой D2D (Device-to-Device), обеспечивающее оптимальные условия для внедрения технологии NOMA. (Напомним, что D2D – это технология связи между двумя мобильными пользователями минуя базовую станцию).

Авторы утверждают, что использование технологии D2D на основе NOMA позволяет эффективно бороться с гетерогенными помехами между различными типами узлов.

В рамках исследуемой концепции передатчик D2D действует не только как ретранслятор пары приемопередатчиков сотовой связи, но и передает свою собственную информацию сопряженному приемнику, с помощью технологии NOMA. Предлагаемый подход повышает надежность и безопасность сетей при наличии устройства перехвата (УП) [33].

Для решения рассматриваемой задачи авторами предлагается протокол совместной работы с оптимизацией мощности. В частности, чтобы предотвратить утечку информации, полнодуплексный (ПДПРМ, full-duplex, FD) приемник сигналов сотовой связи вводит сигналы искусственного шума (AN), с целью ухудшения состояния канала для устройства перехвата, одновременно выполняя формирование лучей диаграммы направленности для подавления AN в направлениях легитимных пользователей. Кроме того, передатчик D2D обеспечивает оптимальные условия по качеству обслуживания с помощью планирования мощности.

Решение задачи совместной оптимизации мощности формулируется путем максимизации коэффициента секретности всей системы в соответствии с требованиями к приемникам сотовой связи и D2D.

Основные результаты, представленные в статье, следующие:

– предлагается безопасный совместный протокол связи D2D на основе NOMA. В протоколе передатчик D2D осуществляет передачу собственного сигнала и ретранслирует групповой сигнал в соответствующие пункты приема. ПДПРМ общей сотовой сети передает сигналы AN и выполняет формирование лучей для противодействия устройству перехвата. Таким образом, создается препятствие для корректной расшифровки, передаваемой между легитимными пользователями информации;

– формулируется общая задача распределения мощности для оптимизации коэффициента секретности совместных сетей с гарантированными требованиями QoS для двух общих сетей;

– для анализа надежности и безопасности структуры сети вводятся параметры вероятности отключения соединения (connection outage probability) и вероятности нарушения секретности системы (secrecy outage probability of the system). Численные и теоретические результаты совпадают с результатами моделирования и свидетельствуют, что совместное применение NOMA с сетями D2D обеспечивает большую безопасность, а также необходимые требования, в сравнении с обычными схемами OMA.

В статье [33] рассматривается общая сеть D2D на основе NOMA, изображенная на рис. 2. В данной сети пользователь сотовой связи «Пользователь 1», намеревается передать конфиденциальные сообщения получателю «Пользователь 2». По причине неудовлетворительных условий передачи (например, затухание энергии, затенение и отсутствие разрешения на доступ) прямой связи между Пользователем 1 и Пользователем 2 нет. Одновременно, пара пользователей сети D2D, состоящая из передатчика (ПРД) и приемника (ПРМ), намеревается осуществить передачу информации в диапазоне сотовых сетей. Для решения данной задачи пользователи D2D занимают спектр сотовой сети, при этом ПРД должен действовать как ретранслятор, чтобы Пользователь 1 мог передать информацию Пользователю 2. Кроме того, устройство перехвата (УП) пытается выполнить перехват информации пары устройств D2D и пользователя сети сотовой связи. Предполагается, что УП находится рядом с приемниками и прослушивает только передачи второго порядка (second hop transmissions). Оборудование Пользователя 2 оснащено N антеннами и работает в полнодуплексном режиме, а другие пользователи с одной антенной работают в полудуплексном режиме.

Рис. 2. Схема общей сети D2D на основе NOMA

Структура передачи информации в сетях разделена на два этапа. В рассматриваемом случае процесс передачи следующий. В рамках первого этапа используется только один канал связи между Пользователем 1 и ПРД, где выполняется передача конфиденциальной информации от Пользователя 1 к ПРД. В рамках второго этапа ПРД декодирует и пересылает ранее полученный сигнал, одновременно излучая собственный сигнал ПРМ посредством кодирования с суперпозицией. Кроме того, чтобы ухудшить канал связи для УП, Пользователь 2 отводит одну антенну для приема сигнала от ПРД, одновременно излучая сигналы AN, чтобы ввести в заблуждение УП, используя оставшиеся N – 1 антенн.

На рисунке 3 показано превосходство предложенной схемы по коэффициенту секретности (КС), по сравнению со альтернативными схемами.

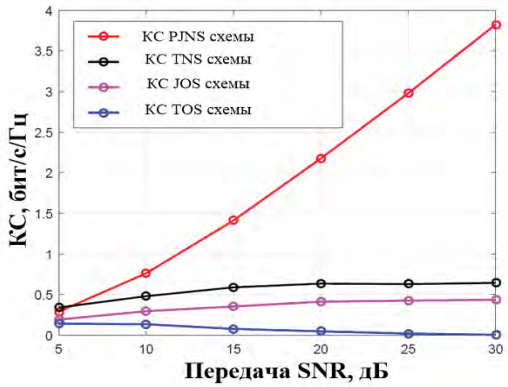


Рис. 3. Зависимость коэффициента секретности (КС) от SNR

Из рисунка 3 видно, что КС альтернативных схем одинаковы, в то время как КС предложенной схемы PJNS быстро растет.

В статье проверено улучшение характеристик безопасности предлагаемой схемы. В частности, предложенная схема снижает скорость прослушивания с помощью сигналов AN, что в конечном итоге вызывает быстрый рост КС системы.

В качестве альтернативных схем для сравнения используются традиционная схема NOMA, схема OMA с помехами и традиционная схема OMA в сетях D2D, которые обозначаются как «TNS», «JOS» и «TOS». Обозначение «PJNS» соответствует предложенной схеме NOMA с использованием AN и коэффициентов распределения мощности.

4. Моделирование системы NOMA

Рассмотрим структурную схему (рис. 4.) системы NOMA с разделением по мощности для случая передачи сигнала четырёх станций к одной базовой станции в режиме пространственного мультиплексирования (SM, Spatial Multiplexing).

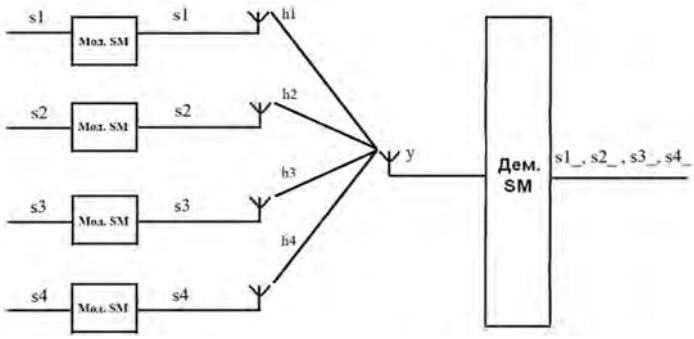


Рис. 4. Структурная схема передачи сигнала четырех станций к одной базовой станции

Для восходящей линии связи, состоящей из четырех абонентских станций, передающих сигнал и принимающей базовой станции, принимаемый базовый сигнал будет иметь вид:

$$y_{\text{БС}} = \sum_{i=1}^{4} \beta_i h_i S_i + n,$$

где β_i – коэффициент распределения мощности для i-го пользователя, h_i – комплексный коэффициент передачи для i-го пользователя, S_i – передаваемый сигнал i-го пользователя, n – омплексный шум в канале связи.

Результаты моделирования приведены на рисунке 5.

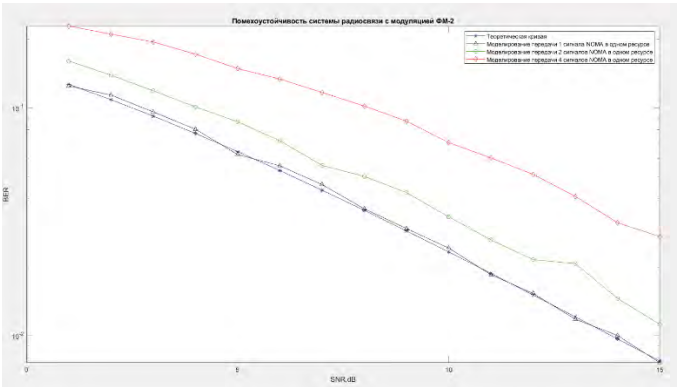


Рис. 5. График зависимости коэффициента ошибок на бит (BER) от SNR при числе испытаний L = 10000

На графике представлены зависимости коэффициента ошибки на бит (BER) от отношения сигнал/шум (SNR) для разного числа сигналов пользователей в одном ресурсе.

Из графика видно, что при числе абонентов не более трех характеристики ухудшаются незначительно по сравнению с OMA (что соответствует 300% загрузки системы).

Алгоритм моделирования приведён в таблице 1.

Таблица 1

Номер шага	Действие программы моделирования	Переменные, используемые в программе
1	Ввод начальных данных (число испытаний, число сигналов)	*L , S*
2	Начало цикла по SNR (отношению сигнал/шум)	*SNR*
	Вычисление среднеквадратического отклонения аддитивного белого гауссовского шума для текущего SNR	*sigma*
3	Начало цикла по L (число испытаний)	*L = 10000*
	Генерирование случайных битов	*b*
	Формирование информационных символов в модуляторах	*s*
4	Генерирование значения шума с заданным среднеквадратическим отклонением и коэффициентов передачи	*n, h*
5	Получение смеси сигналов и шума	$y = \sum_{i}^{1} h_i s_i + n$
6	Демодуляция (формирование оценки символа)	*s_*
7	Формирование оценки принятого бита	*b_*
8	Определение факта битовой ошибки при приёме (сравнение величин b и b)	*err*
9	Подсчёт общего числа ошибок	*sum*
10	Завершение цикла по пункту 3	
11	Вычисление коэффициента битовых ошибок	*BER*
12	Получение теоретической кривой для релеевского канала	*BERtheoryR*
13	Завершение цикла по пункту 2	
14	Построение теоретической кривой	*BERtheoryR*
15	Построение кривых коэффициента ошибок на бит для разного числа абонентов	*BER*

Заключение

В этой работе был проведен анализ решений повышения секретности систем беспроводной связи, использующих технологию NOMA в разных сценариях, в том числе сценарий D2D и безопасная совместная система связи на основе IoT, в основном по зарубежным публикациям. В когнитивных радиосетях CRN с поддержкой NOMA, применяя алгоритм совместной оптимизации в стратегии безопасной передачи, можно получить допустимое решение для распределения мощности, чтобы максимизировать коэффициент секретности SSR всех пользователей. Для стратегии совместной безопасной передачи с ретрансляцией в системах IoT было рассмотрено два вида схем.

В первом случае рассматриваются пассивные подслушиватели, для этой схемы предложено оптимизировать, для безопасной передачи, векторы формирования лучей с помощью искусственного шума. Во втором случае рассматриваются активные подслушиватели, такой способ предполагает использование лучей на основе ортогональной проекции. В обоих случаях благодаря системе NOMA спектральная эффективность и секретность растёт. В сценарии D2D система NOMA позволяет увеличить надёжность и безопасность сети, при наличии устройства перехвата, благодаря не только ретрансляции, а также передачи информации с передатчика D2D сопряженному приемнику. При этом характеристики помехоустойчивости с увеличением абонентов NOMA ухудшаются незначительно и находятся в допустимых пределах.

Литература

1. *L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I, and Z. Wang*. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends // IEEE Commun. Mag., vol. 53, pp. 74-81, Sept. 2015.

2. *Q.C. Li, H. Niu, A.T. Papathanassiou, G. Wu.* 5G network capacity: Key elements and technologies // IEEE Veh. Technol. Mag., vol. 9, pp. 71-78, Mar. 2014.

3. *L. Wei, T. Jing, X. Fan, Y. Wen, Y. Huo*. The Secrecy Analysis over Physical Layer in NOMA-enabled Cognitive Radio Networks // IEEE, 978-1-5386-3180-5/18, 2018.

4. *Z. Li, T. Jing, X. Cheng, Y. Huo, W. Zhou, D. Chen*. Cooperative jamming for secure communications in MIMO cooperative cognitive radio networks // 2015 IEEE International Conference on Communications (ICC), June 2015, pp. 7609-7614.

5. *P.H. Lin, F. Gabry, R. Thobaben, E.A. Jorswieck, M. Skoglund*. Multi-phase smart relaying and cooperative jamming in secure cognitive radio networks // IEEE Transactions on Cognitive Communications and Networking, vol. 2, no. 1, pp. 38-52, March 2016.

6. *Y.Y. He, J. Evans, S. Dey*. Secrecy rate maximization for cooperative overlay cognitive radio networks with artificial noise // IEEE International Conference on Communications, 2014, pp. 1663-1668.

7. *X. Xing, T. Jing, W. Cheng, Y. Huo, X. Cheng, T. Znati.* Cooperative spectrum prediction in multi-PU multi-SU cognitive radio networks // Mobile Networks & Applications, vol. 19, no. 4, pp. 502-511, 2014.

8. *X. Xing, T. Jing, W. Cheng, Y. Huo, X. Cheng*. Spectrum prediction in cognitive radio networks // IEEE Wireless Communications, vol. 20, no. 2, pp. 90-96, April 2013.

9. *Y. Pei, Y.C. Liang, K.C. Teh, K.H. Li.* Secure communication in multiantenna cognitive radio networks with imperfect channel state information // IEEE Transactions on Signal Processing, vol. 59, no. 4, pp. 1683-1693, April 2011.

10. *Y. Liu, Z. Ding, M. Elkashlan, H.V. Poor*. Cooperative nonorthogonal multiple access with simultaneous wireless information and power transfer // IEEE Journal on Selected Areas in Communications, vol. 34, no. 4, pp. 938-953, 2016.

11. *K. Jiang, T. Jing, Z. Li, Y. Huo, F. Zhang*. Analysis of secrecy performance in fading multiple access wiretap channel with SIC receiver // IEEE Conference on Computer Communications, May 2017, pp. 1-9.

12. *Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, H.V. Poor*. Application of non-orthogonal multiple access in LTE and 5G networks // IEEE Communications Magazine, vol. 55, no. 2, pp. 185-191, 2015.

13. *Z. Ding, P. Fan, H.V. Poor*. Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions // IEEE Transactions on Vehicular Technology, vol. 65, no. 8, pp. 6010-6023, 2016.

14. *N. Zabetian, M. Baghani, A. Mohammadi*. Rate optimization in NOMA cognitive radio networks," in International Symposium on Telecommunications, 2017, pp. 62-65.

15. *L. Yang, J. Chen, Q. Ni, J. Shi, X. Xue*. NOMA-enabled cooperative unicast-multicast: Design and outage analysis // IEEE Transactions on Wireless Communications, vol. 16, no. 12, pp. 7870-7889, Dec 2017.

16. *Z. Qin, Y. Liu, Z. Ding, Y. Gao, M. Elkashlan*. Physical layer security for 5G non-orthogonal multiple access in large-scale networks // 2016 IEEE International Conference on Communications (ICC), May 2016, pp. 1-6.

17. *Y. Zhang, H.M. Wang, Q. Yang, Z. Ding*. Secrecy sum rate maximization in non-orthogonal multiple access // IEEE Communications Letters, vol. 20, no. 5, pp. 930-933, 2016.

18. *Y. Liu, Z. Ding, M. Elkashlan, J. Yuan.* Nonorthogonal multiple access in large-scale underlay cognitive radio networks // IEEE Transactions on Vehicular Technology, vol. 65, no. 12, pp. 10 152-10 157, Dec 2016.

19. *L. Hu, H. Wen, B. Wu, F. Pan, R. F. Liao, H. Song, J. Tang, X. Wang.* Cooperative jamming for physical layer security enhancement in Internet of Things // IEEE Internet of Things Journal, vol. 5, no. 1, pp. 219-228, Feb. 2018.

20. *M. Tian, Q. Zhang, S. Zhao, Q. Li, J. Qin*. Secrecy sum rate optimization for downlink MIMO nonorthogonal multiple access systems // IEEE Signal Process. Lett., vol. 24, no. 8, pp. 1113-1117, Aug. 2017.

21. *M.F. Hanif, Z. Ding, T. Ratnarajah, G.K. Karagiannidis*. A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems // IEEE Trans. Signal Process., vol. 64, no. 1, pp. 76-88, Jan. 2016.

22. Q. *Sun, S. Han, C. L. I, Z. Pan*. On the ergodic capacity of MIMO NOMA systems // IEEE Wireless Commun. Lett., vol. 4, no. 4, pp. 405-408, Aug. 2015.

23. *Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, Z. Zhong.* Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems // IEEE Trans. Signal Process., vol. 65, no. 18, pp. 4874-4886, Sept. 2017.

24. *M. Zeng, A. Yadav, O.A. Dobre, G.I. Tsiropoulos, H.V. Poor*. Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster // IEEE J. Sel. Areas in Commun., vol. 35, no. 10, pp. 2413-2424, Oct. 2017.

25. *S.L. Keoh, S.S. Kumar, H. Tschofenig.* Securing the internet of things: A standardization perspective // IEEE Internet Things Journal, vol. 1, no. 3, pp. 265-275, Jun. 2014.

26. *J. Granjal, E. Monteiro, J.S. Silva.* Security for the internet of things: A survey of existing protocols and open research issues // IEEE Commun. Surveys Tuts., vol. 17, no. 3, pp. 1294-1312, 3rd Quart., 2015.

27. *Y.W.P. Hong, P.C.Lan, and C.C.J. Kuo.* Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches // IEEE Signal Process. Mag., vol. 30, no. 5, pp. 29-40, Sept. 2013.

28. *L. Hu, H. Wen, B. Wu, J. Tang, F. Pan*. Adaptive secure transmission for physical layer security in cooperative wireless networks // IEEE Commun. Lett., vol. 21, no. 3, pp. 524-527, Mar. 2017.

29. *A. Behnad, M.B. Shahbaz, T.J. Willink, X. Wang.* Statistical analysis and minimization of security vulnerability region in amplify and-forward cooperative systems // IEEE Trans. Wireless Commun., vol. 16, no. 4, pp. 2534-2547, Apr. 2017.

30. *X. Hu, P. Mu, B. Wang, Z. Li.* On the secrecy rate maximization with uncoordinated cooperative jamming by single-antenna helpers // IEEE Trans Veh. Technol., vol. 66, no. 5, pp. 4457-4462, May 2017.

31. *Y. Zhang, Y. Shen, H. Wang, J. Yong, X. Jiang*. On secure wireless communications for IoT under eavesdropper collusion // IEEE Trans. Autom. Sci. Eng., vol. 13, no. 3, pp. 1281-1293, July 2016.

32. *P. Huang, Y. Hao, T. Lv, J. Xing, P. Takis*. Secure Beamforming Design in Relay-Assisted Internet of Things // IEEE. Autom. Sci. Eng, 2327-4662, 2018.

33. *Q. Li, P. Ren, D. Xu.* Security Enhancement and QoS Provisioning for NOMA-Based Cooperative D2D Networks // IEEE Access, vol. 7, pp. 129387-129401, 2019, doi: 10.1109/ACCESS.2019.2939783.

34. *Панкратов Д.Ю., Кумецкий П.Ю., Маков М.В.* Неортогональный множественный доступ с разделением по мощности для восходящей и нисходящей линии связи. 2022. № 4. С. 1-6.

35. *Филатова Е.Е., Панкратов Д.Ю.* Моделирование передачи в системе NOMA с разным числом станций. Информационная безопасность NOMA. 2022. № 3. С. 1-7.

SECURITY ANALYSIS OF NOMA SYSTEMS

SERGEY S. ALEKSEEV,
Moscow, Russia, saleks00@yandex.ru

TATYANA P. KOSICHKINA,
Moscow, Russia, t.p.kosichkina@mtuci.ru

DENIS YU. PANKRATOV,
Moscow, Russia, dpankr@mail.ru

ILYA A. SHAMSUTDINOV,
Moscow, Russia, malon228@mail.ru

KEYWORDS: *NOMA, security, PLS, MIMO-NOMA, IoT, D2D.*

ABSTRACT

Introduction: Existing Physical Layer Security (PLS) schemes for NOMA are either based on cryptography-based approaches or are limited to approaches that require high processing complexity and key sharing. The conventional NOMA system suffers from security risks and weaknesses, such as being prone to external or internal eavesdropping. By sending NOMA messages to multiple users at the same time on the same resources, there is a risk that an unauthorized user could eavesdrop on or access multiple users' information if the NOMA transmission is successfully intercepted. The NOMA system is susceptible to internal interception when protecting confidential data in cases where untrusted users are present. Therefore, traditional methods of data exchange cannot provide the required level of secrecy for NOMA systems. **The results.** The work uses non-orthogonal access (NOMA) technology and security assessment of various wireless communication systems using NOMA, such as cognitive radio networks, systems with support for relaying the IoT mode, as well as the D2D mode. The results of the analysis of security characteristics and operation of the NOMA system in different modes are presented.

REFERENCES

1. L. Dai, B. Wang, Y. Yuan, S. Han, C. I. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag*., vol. 53, pp. 74-81, Sept. 2015.

2. Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag*., vol. 9, pp. 71-78, Mar. 2014.

3. L. Wei, T. Jing, X. Fan, Y. Wen, Y. Huo, "The Secrecy Analysis over Physical Layer in NOMA-enabled Cognitive Radio Networks", *IEEE*, 978-1-5386-3180-5/18, 2018.

4. Z. Li, T. Jing, X. Cheng, Y. Huo, W. Zhou, and D. Chen, "Cooperative jamming for secure communications in MIMO cooperative cognitive radio networks," *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 7609-7614.

5. P. H. Lin, F. Gabry, R. Thobaben, E. A. Jorswieck, and M. Skoglund, "Multi-phase smart relaying and cooperative jamming in secure cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 1, pp. 38-52, March 2016.

6. Y. Y. He, J. Evans, and S. Dey, "Secrecy rate maximization for cooperative overlay cognitive radio networks with artificial noise," *IEEE International Conference on Communications*, 2014, pp. 1663- 1668.

7. X. Xing, T. Jing, W. Cheng, Y. Huo, X. Cheng, and T. Znati, "Cooperative spectrum prediction in multi-PU multi-SU cognitive radio networks," *Mobile Networks & Applications*, vol. 19, no. 4, pp. 502-511, 2014.

8. X. Xing, T. Jing, W. Cheng, Y. Huo, and X. Cheng, "Spectrum prediction in cognitive radio networks," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 90-96, April 2013.

9. Y. Pei, Y. C. Liang, K. C. Teh, and K. H. Li, "Secure communication in multiantenna cognitive radio networks with imperfect channel state information," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1683-1693, April 2011.

10. Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative nonorthogonal multiple access with simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 938-953, 2016.

11. K. Jiang, T. Jing, Z. Li, Y. Huo, and F. Zhang, "Analysis of secrecy performance in fading multiple access wiretap channel with SIC receiver," *IEEE Conference on Computer Communications*, May 2017, pp. 1-9.

12. Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185-191, 2015.

13. Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010-6023, 2016.

14. N. Zabetian, M. Baghani, and A. Mohammadi, "Rate optimization in NOMA cognitive radio networks," *International Symposium on Telecommunications*, 2017, pp. 62-65.

15. L. Yang, J. Chen, Q. Ni, J. Shi, and X. Xue, "NOMA-enabled cooperative unicast-multicast: Design and outage analysis," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 7870-7889, Dec 2017.

16. Z. Qin, Y. Liu, Z. Ding, Y. Gao, and M. Elkashlan, "Physical layer security for 5G non-orthogonal multiple access in large-scale networks," *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1-6.

17. Y. Zhang, H. M. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Communications Letters,* vol. 20, no. 5, pp. 930-933, 2016.

18. Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Nonorthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 152-10 157, Dec 2016.

19. L. Hu, H. Wen, B. Wu, F. Pan, R. F. Liao, H. Song, J. Tang, and X. Wang, "Cooperative jamming for physical layer security enhancement in Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 219-228, Feb. 2018.

20. M. Tian, Q. Zhang, S. Zhao, Q. Li, and J. Qin, "Secrecy sum rate optimization for downlink MIMO nonorthogonal multiple access systems," *IEEE Signal Process. Lett*., vol. 24, no. 8, pp. 1113-1117, Aug. 2017.

21. M.F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76-88, Jan. 2016.

22. Q. Sun, S. Han, C. L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405- 408, Aug. 2015.

23. Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process*., vol. 65, no. 18, pp. 4874-4886, Sept. 2017.

24. M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, "Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster," *IEEE J. Sel. Areas in Commun*., vol. 35, no. 10, pp. 2413-2424, Oct. 2017.

25. S. L. Keoh, S. S. Kumar, and H. Tschofenig, "Securing the internet of things: A standardization perspective," *IEEE Internet Things Journal*, vol. 1, no. 3, pp. 265-275, Jun. 2014.

26. J. Granjal, E. Monteiro, and J. S. Silva, "Security for the internet of things: A survey of existing protocols and open research issues," *IEEE Commun. Surveys Tuts.,* vol. 17, no. 3, pp. 1294-1312, 3rd Quart., 2015.

27. Y. W. P. Hong, P. C. Lan, and C. C. J. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Process. Mag.,* vol. 30, no. 5, pp. 29-40, Sept. 2013.

28. L. Hu, H. Wen, B. Wu, J. Tang, and F. Pan, "Adaptive secure transmission for physical layer security in cooperative wireless networks," I*EEE Commun. Lett.*, vol. 21, no. 3, pp. 524-527, Mar. 2017.

29. A. Behnad, M. B. Shahbaz, T. J. Willink, and X. Wang, "Statistical analysis and minimization of security vulnerability region in amplify and-forward cooperative systems," *IEEE Trans. Wireless Commun*., vol. 16, no. 4, pp. 2534-2547, Apr. 2017.

30. X. Hu, P. Mu, B. Wang, and Z. Li, "On the secrecy rate maximization with uncoordinated cooperative jamming by single-antenna helpers," *IEEE Trans Veh. Technol*., vol. 66, no. 5, pp. 4457-4462, May 2017.

31. Y. Zhang, Y. Shen, H. Wang, J. Yong, and X. Jiang, "On secure wireless communications for IoT under eavesdropper collusion," *IEEE Trans. Autom. Sci. Eng*., vol. 13, no. 3, pp. 1281-1293, July 2016.

32. P. Huang, Y. Hao, T. Lv, J. Xing, and P. Takis, "Secure Beamforming Design in Relay-Assisted Internet of Things", *IEEE. Autom. Sci. Eng,* pp. 2327-4662, 2018.

33. Q. Li, P. Ren and D. Xu, "Security Enhancement and QoS Provisioning for NOMA-Based Cooperative D2D Networks", *IEEE Access*, vol. 7, pp. 129387-129401, 2019, doi: 10.1109/ACCESS.2019.2939783.

34. D.Yu. Pankratov, P.Yu. Kumetsky, M.V. Makov, “Non-orthogonal power division multiple access for uplink and downlink,” 2022. No. 4, pp. 1-6.35.

35. E.E. Filatova, D.Yu. Pankratov, “Simulation of transmission in the NOMA system with different numbers of stations,” *Information security NOMA*. 2022. No. 3, pp. 1-7.

INFORMATION ABOUT AUTHORS:

Sergey S. Alekseev, MTUCI, student of ZRS1701, Moscow Technical University of Communications and Informatics, Moscow, Russia
Tatyana P. Kosichkina, MTUCI, Ph.D., Associate Professor, cafe of SSSRT, Moscow, Russia
Denis Yu. Pankratov, MTUCI, Ph.D., Associate Professor, cafe of SSSRT, Moscow Technical University of Communications and Informatics, Moscow, Russia
Ilya A. Shamsutdinov, MTUCI, student of MIT2201, Moscow Technical University of Communications and Informatics, Moscow, Russia

For citation: Alekseev S.S., Kosichkina T.P., Pankratov D.Yu., Shamsutdinov I.A. Security analysis of NOMA systems. H&ES Reserch. 2023. Vol. 15. No 4. P. 38-46. doi: 10.36724/2409-5419-2023-15-4-38-46 (In Rus)

doi: 10.36724/2409-5419-2023-15-4-47-59

ТРЕТЬЯ ПЛАТФОРМА ИНФОРМАТИЗАЦИИ И BIG DATA

МОШАК
Николай Николаевич[1]

РУДИНСКАЯ
Сабина Романовна[2]

ГРУЗДЕВ
Алексей Андреевич[3]

АННОТАЦИЯ

Введение: несмотря на то, что внедрение Больших данных во многие сферы жизни является неизбежной реальностью ближайшего будущего, имеются сложности с классификацией информации, выделением знаний, хранением, вопросами безопасности и масштабируемости. В работе описывается понятие третьей платформы информатизации и Больших Данных, играющих важную роль в этом процессе. Поставлены проблемы, связанные с необходимостью эффективной обработки и анализа Больших Данных. Показаны и рассмотрены сферы жизни, в которые активно внедряются и развиваются новые технологии. **Целью исследования является** обсуждение вопросов трендов внедрения Больших Данных в рамках третьей платформы информатизации. Исследуются проблемы Больших Данных и некоторые пути их решения. Рассматриваются взаимосвязи и закономерности в области Больших Данных и их влияние на развитие информационных технологий. **Методика проведения исследования:** используются аналитические методы исследования, основанные на анализе данных и концепций третьей платформы информатизации. **Результаты исследования:** проведен глубокий анализ для полного погружения в вопрос третьей платформы информатизации и понятия Больших Данных. Показаны преимущества внедрения компонентов технологий в сферы жизни, а также изучены проблемы Больших Данных и некоторые пути их решения. Приводятся основные теоретические результаты исследования, подтверждающие взаимосвязи и закономерности между элементами развивающихся информационных технологий. Обсуждается практическая значимость третьей платформы информатизации и ее компонентов. Предлагаются рекомендации, связанные с применением этих технологий в реальных условиях, а также их взаимное влияние технологий для достижения более эффективных результатов.

Сведения об авторах:

[1] д.т.н., профессор, профессор Санкт-Петербургского государственного университета аэрокосмического приборостроения; профессор Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, г. Санкт-Петербург, Россия, nnmoshak49@mail.ru

[2] Старший преподаватель Учреждения образования "Белорусская государственная академия связи", Минск, Республика Беларусь, sabina.rudin@mail.ru

[3] Студент Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, г. Санкт-Петербург, Россия, gruzdev.a.a26@mail.ru

КЛЮЧЕВЫЕ СЛОВА: *анализ информации, Big Data, виртуализация, кластеризация, обработка данных, третья платформа информатизации.*

Для цитирования: Мошак Н.М., Рудинская С.Р., Груздев А.А. Третья платформа информации и Big Data // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 47-59. doi: 10.36724/2409-5419-2023-15-4-47-59

Введение

Конец 20-го и начало 21 века характеризуются большим скачком в увеличении объема использования цифровой информации. При этом задача структуризации, обработки, выделения закономерностей и внедрения в производственные процессы «Больших Данных» (Big Data с англ. Большие Данные) становится наиболее актуальной [1].

В сущности, понятие Big Data подразумевает работу с информацией огромного объёма и разнообразного состава, которая часто обновляется и может располагаться в различных информационных источниках. Актуальность использования технологии Big Data достаточно велика, так как в данный момент она является одним из ключевых драйверов развития информационных технологий. Это направление получило широкое распространение в западных странах и в России. В исследованиях, связанных с Большими Данными, изучаются проблемы и задачи обработки и их хранения, а также способы выделения знаний из них. Они ведутся в рамках различных дисциплин и областей, таких как информационные науки, моделирование неопределённости, машинное обучение, статистическое обучение, распознавание образов, методы хранения данных, обработка сигналов и т.д.

В статье обсуждаются вопросы внедрения и проблемы Больших Данных, связанные с обработкой, хранением и анализом, выделением знаний и защитой Больших Данных, в рамках третьей платформы информатизации.

Необходимость внедрения и применения Больших Данных в рамках третьей платформы информатизации

Термин Big Data был предложен редактором журнала Nature Клиффорд Линч в 2008 году, в сентябрьском спецвыпуске «Как могут повлиять на будущее науки технологии, открывающие возможности работы с большими объёмами данных?». В нём говорилось о феномене взрывного роста объёмов и многообразия обрабатываемых данных в мире, а также о технологических перспективах в решении задачи их обработки [2]. Согласно [3, 4] термин Big Data относится к наборам данных, размер которых превосходит возможности типовых баз данных (БД) по хранению, управлению и анализу. Клиффорд Линч отнес к Big Data любые массивы неоднородных данных, превышающие объёмом обработки 150 Гб в сутки. Под этим понятием понимают структурированные или неструктурированные (в большинстве своём) массивы данных большого объема. Однако единого критерия на объём до сих пор не существует в силу особенностей данных, получаемых из разных источников. Вот только несколько примеров источников таких объемов [3]:

1) Нью-Йоркская фондовая биржа генерирует около терабайта данных в день.

2) Об м хранилища социальной сети Facebook каждый день увеличивается на 500 терабайт.

3) Проект Internet Archive уже хранит 2 петабайта данных и прирастает 20 терабайтами в месяц.

4) Эксперименты на Большом адронном коллайдере могут генерировать около петабайта данных в секунду.

Компания Gartner определяет Большие Данные как большие объемы, высокоскоростные и/или разнообразные информационные активы, требующие рентабельных, инновационных форм обработки информации, которые обеспечивают более глубокое понимание, принятие решений и автоматизацию процессов [5]. Большие Данные классифицируются в зависимости от типа источника данных (Интернет или медиа-источники, сгенерированные машинные данные), по формату контента (структурированные, полуструктурированные, неструктурированные), по способу хранения данных и т.д. Не структурированные наборы данных, имея разный вид, особенности заполнения, поля, источники – не имеют единого вида или общего формата, из-за чего становится сложно анализировать их единым механизмом. Храниться они могут в виде текста, электронных писем, видео, пакетов или мультимедийных файлов. Слабоструктурированные данные представляет собой некий гибрид или смешанную категорию между структурированными и неструктурированными данными. Основное отличие слабоструктурированных данных заключается в том, что их нельзя категорировать, но они имеют некоторые определенные свойства (например, логи, тэги), которые можно проанализировать и структурировать для хранения. Для наглядности на рисунке 1 представлена общая классификация технологии Больших Данных [6].

Рис. 1. Общая классификация технологии Больших данных

Основные компоненты третьей платформы информатизации как «платформы Больших Данных»

С появлением Больших Данных и необходимостью их эффективной обработки, важную роль играет третья платформа информатизации. Третья платформа (the 3rd platform, the Third platform – англ.) – термин, предложенный аналитиками IDC для описания новой парадигмы развития компьютерной индустрии [3]. По мнению экспертов компании, «первой» платформой компьютерной эры были мейнфреймы, мощные серверы, которые позволили широко оценить мощь и возможности компьютерных вычислений, «вторая» – на базе ПК и клиент-серверной архитектуры с сотнями тысяч приложений и сотнями миллионов пользователей.

Третья платформа информатизации охватывает целый спектр современных технологий, составляющими которой являются облачные технологии, технологии Интернета

вещей, Больших Данных, мобильного широкополосного доступа, сервисы Over The Top (OTT-сервисы, наложенные сервисы), облачные технологии, «умная экономика», соцсети – с миллиардами пользователей (рис. 2) [7]. Технологии третьей платформы информатизации позволяют создавать информационные инфраструктуры различного уровня сложности и назначения благодаря конвергенции технологий хранения, передачи и обработки данных. Третья платформа, также известная как «платформа Больших Данных», объединяет несколько технологических компонентов и концепций, которые обеспечивают возможности работы с данными масштаба Больших Данных.

Причины, по которым обозначенные четыре составляющие (облако, мобильность, соцсети и «Большие Данные») были объединены в одну платформу, представлены в докладе менеджера по исследованиям IDC Александра Прохорова на конференции IDC по управлению контентом (рис. 3) [8]. Взаимосвязь указанных технологий можно проследить не только по горизонтали, как это указано на рисунке, но и в любых сочетаниях.

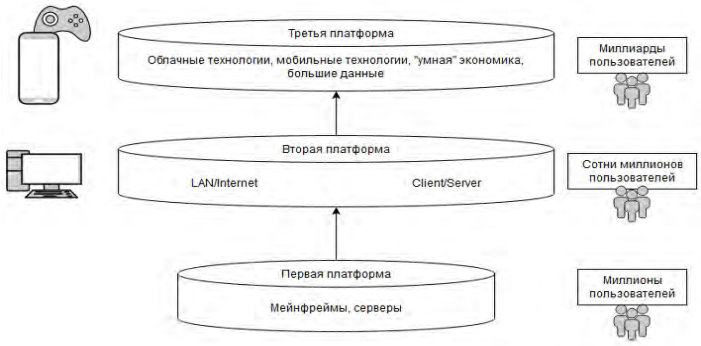


Рис. 2. «Третья платформа» IDC – очередной этап развития ИТ

Рис. 3. Взаимное влияние технологий, составляющих понятие «третья платформа»

Действительно, пользователи растущего числа мобильных устройств производят всё больше контента, который удобно хранить в облаках. Контент, помещенный в облако, может «раздаваться» владельцам различных мобильных платформ в соответствующих форматах. Количество мобильных устройств непрерывно растет, поэтому потребность в облачном хранилище постоянно увеличивается. Вот несколько ключевых аспектов роли третьей платформы информатизации:

1) *Обеспечение масштабируемости*. Третья платформа информатизации предоставляет масштабируемые архитектуры и технологии, такие как распределенные системы обработки данных и облачные вычисления, которые позволяют масштабировать инфраструктуру и обрабатывать данные в параллельном режиме.

2) *Использование открытых стандартов и технологий*. Третья платформа информатизации активно использует открытые стандарты и технологии, что способствует унификации и совместимости различных систем обработки данных. Это позволяет интегрировать различные источники данных и инструменты анализа, обеспечивая более эффективную работу с Большими Данными.

3) *Использование аналитики больших данных*. Третья платформа информатизации предоставляет инструменты и методы для анализа Больших Данных. Это включает в себя технологии добычи знаний (Data Mining), машинное обучение, статистический анализ и другие подходы, которые позволяют находить закономерности, тенденции и паттерны в больших объемах информации. Аналитика Больших Данных помогает превратить необработанные данные в ценные знания и понимание для принятия решений.

4) *Обеспечение реального времени*. Третья платформа информатизации позволяет обрабатывать и анализировать данные в реальном времени. Это важно для оперативного реагирования на изменения и события, основанных на данных. Реально временная обработка данных позволяет организациям принимать более точные и своевременные решения, основанные на актуальной информации.

5) *Управление и хранение данных*. Третья платформа информатизации предоставляет средства для управления и хранения больших объемов данных. Это включает в себя разработку и использование распределенных баз данных, систем управления данными, технологии хранения данных, такие как Hadoop и NoSQL. Они обеспечивают эффективное хранение, организацию и доступ к данным большого объема.

6) *Интеграция различных источников данных*. Третья платформа информатизации способствует интеграции и объединению данных из различных источников. Это включает в себя структурированные и неструктурированные данные, данные из социальных сетей, мобильных устройств, сенсоров и других источников. Интеграция данных позволяет создавать более полную и всестороннюю картину, а также обогащать и улучшать аналитические модели и выводы.

Роль третьей платформы информатизации в работе с Большими Данными состоит в обеспечении необходимых технологий, инструментов и методов для эффективной обработки, анализа и использования данных масштаба Больших Данных. Она позволяет организациям извлекать ценные знания, улучшать принятие решений, повышать конкурентоспособность и создавать новые возможности для развития и инноваций. Решение данных проблем требует разработки и использования специализированных технологий, таких как распределенные базы данных, облачные хранилища, технологии обработки потоков данных и многопоточных алгоритмов. Однако подходы к решению этих проблем должны варьироваться в зависимости от конкретного контекста и требований организации или проекта [9].

Сферы внедрения и применения Больших Данных

Технологии Больших Данных дают возможность обработать огромный объем неструктурированных данных, проанализировать их, систематизировать и выявить закономерности там, где человек их, не смог бы обнаружить [4].

Следует отметить, что объемы обрабатываемых Больших Данных непрерывно увеличиваются, как и увеличивается скорость их обработки. Большинство крупнейших компаний – поставщиков информационных технологий (IBM, Oracle, Microsoft, Hewlett-Packard, EMC), для организации рабочего процесса начинают использовать технологию Больших Данных. Так, компания Gartner [5] отмечает Большие Данные как тренд номер два в информационно-технологической инфраструктуре (после виртуализации), а также прогнозирует, что внедрение данной технологии окажет наибольшее влияние на информационные технологии в производстве, торговле, здравоохранении, государственном управлении и других в сферах и отраслях, где регистрируются частые перемещения информационных ресурсов. Другими словами, для данных сфер задача структуризации, обработки, выделения закономерностей и внедрения в производственные процессы большого объёма пользовательской информации наиболее актуальна.

В качестве примера использования Больших Данных можно привести множество сфер человеческой жизни, которые уже сложно представить без применения этой технологии. Вот некоторые из основных применений больших данных.

Здравоохранение и медицина. Применение Больших Данных в здравоохранении и медицине имеет огромный потенциал для совершенствования системы здравоохранения, улучшения качества жизни пациентов и развития медицинской науки и практики.

1) *Индивидуальное лечение и персонализированная медицина*. Большие Данные позволяют анализировать медицинскую информацию о больших группах пациентов и выявлять паттерны и тренды, связанные с различными заболеваниями. Это позволяет разрабатывать персонализированные лечебные протоколы, предсказывать риски заболеваний и принимать решения о наилучшем лечении для каждого пациента. В медицине Большие Данные помогают с анализом медицинских записей и статистики использования лекарств, диагностикой заболеваний, разработкой лекарств, оптимизацией системы здравоохранения и предоставлением персонализированного медицинского ухода. Анализ и использование Больших Данных в этой области способствует повышению эффективности предоставления услуг.

2) *Раннее выявление заболеваний*. Анализ Больших Данных позволяет выявлять ранние признаки заболеваний и прогнозировать их развитие. Путем мониторинга медицинских данных, включая симптомы, биометрические показатели и результаты тестов, можно обнаружить заболевания на ранних стадиях и принять меры по их предотвращению или раннему лечению.

3) *Улучшение диагностики и обработки изображений*. Большие Данные позволяют анализировать медицинские изображения, такие как рентгены, МРТ, КТ и другие, с использованием алгоритмов машинного обучения и искусственного интеллекта. Это позволяет повысить точность диагностики, обнаружить скрытые патологии и помочь врачам принимать более информированные решения о лечении.

4) *Прогнозирование эпидемий и общественного здоровья*. Анализ Больших Данных позволяет прогнозировать распространение эпидемий и мониторить общественное здоровье. Использование данных о заболеваемости, симптомах, контактах, местоположении и других факторах позволяет предсказывать и контролировать распространение инфекционных заболеваний, а также принимать меры по обеспечению безопасности и здоровья населения.

5) *Управление медицинскими ресурсами*. Большие Данные помогают оптимизировать использование медицинских ресурсов, таких как госпитали, лекарства и медицинское оборудование. Анализ данных позволяет прогнозировать спрос на медицинские услуги, оптимизировать распределение ресурсов, улучшить планирование и принимать решения о рациональном использовании медицинских ресурсов.

Транспорт и логистика. Большие Данные используются для оптимизации маршрутов, управления транспортными потоками, прогнозирования спроса на транспортные услуги и повышения эффективности логистических операций. Вот некоторые области, где Большие Данные находят применение в транспорте и логистике.

1) *Прогнозирование и оптимизация транспортного спроса*. Большие данные позволяют анализировать накопленные данные о перевозках, потоках пассажиров и грузов, транспортных расписаниях и других параметрах, чтобы прогнозировать спрос на транспорт и разрабатывать оптимальные маршруты и графики движения. Это помогает снизить заторы, улучшить планирование маршрутов и повысить эффективность использования транспортных ресурсов.

2) *Управление логистическими сетями*. Большие Данные используются для управления сложными логистическими сетями, включая склады, центры распределения, транспортные маршруты и процессы доставки. Анализ данных позволяет оптимизировать складские запасы, прогнозировать потребности в запасных частях и материалах, улучшать планирование доставки и повышать точность отслеживания грузов.

3) *Безопасность и мониторинг транспорта*. Большие Данные играют важную роль в обеспечении безопасности транспортных систем. С помощью анализа данных о движении транспортных средств, датчиков, видеокамер и других источников можно выявлять аномалии, предсказывать потенциальные аварийные ситуации и предпринимать меры по предотвращению несчастных случаев на дорогах. Большие Данные также используются для мониторинга экологических показателей, энергоэффективности и соблюдения стандартов в транспортной отрасли.

4) *Умные транспортные системы*. Большие Данные способствуют развитию умных транспортных систем, которые используют передовые технологии, такие как датчики, интернет вещей и искусственный интеллект, для улучшения управления транспортом и обеспечения комфортного и безопасного передвижения. Примеры включают системы управления светофорами на основе данных о потоках транспорта, предупреждающие системы о преградах и аварийных ситуациях, а также системы сбора платы за проезд на основе данных о проезжающих транспортных средствах.

5) *Анализ и оптимизация энергопотребления*. Большие Данные используются для анализа энергопотребления в транспортной отрасли и разработки эффективных методов снижения потребления топлива и выбросов вредных веществ. Анализ данных позволяет оптимизировать маршруты, улучшать экономию топлива, внедрять электромобили и альтернативные источники энергии, а также разрабатывать стратегии по сокращению углеродного следа транспорта.

Применение Больших Данных в транспорте и логистике открывает новые возможности для оптимизации процессов, повышения эффективности и улучшения условий перевозок. Однако, необходимо учитывать вопросы конфиденциальности и защиты данных, а также обеспечивать этичное использование информации, связанной с пассажирами и грузами.

Научные исследования. Применение Больших Данных в науке и исследованиях играет важную роль в различных дисциплинах, позволяя ученым обрабатывать и анализировать огромные объемы данных для получения новых знаний, прогнозирования тенденций, идентификации паттернов и развития новых моделей и теорий [10,11]. Вот некоторые области, где Большие Данные имеют значительное влияние на науку и исследования.

1) *Геномика и биологические исследования*. Большие Данные в геномике позволяют исследователям анализировать огромное количество генетической информации для выявления генетических вариантов, связанных с заболеваниями и разработки новых подходов к лечению заболеваний. Анализ геномных данных позволяет выявить генетические мутации, которые могут быть связаны с различными заболеваниями.

2) *Астрономия и космические исследования*. Современные телескопы и спутники генерируют огромные объемы данных о Космосе. Большие Данные используются для анализа данных со спутников, телескопов и других приборов для изучения космических явлений, таких как черные дыры, галактики и звезды, планеты и другие космические объекты, исследовать эволюцию Вселенной, искать экзопланеты и анализировать космические явления.

3) *Экологические исследования*. Большие Данные позволяют ученым анализировать информацию о состоянии окружающей среды, климатических изменениях, распределении видов и изменении экосистем. Анализ данных помогает идентифицировать тренды, прогнозировать последствия климатических изменений, разрабатывать меры по охране окружающей среды и предотвращению угроз биоразнообразию.

4) *Математические моделирование и симуляции*. Большие Данные используются для создания математических моделей и симуляций в различных научных областях. Это позволяет ученым проводить эксперименты, изучать сложные системы, прогнозировать поведение и оценивать эффективность различных стратегий и политик.

Применение Больших Данных в науке и исследованиях обеспечивает новые возможности для открытий и инноваций в различных областях знания. Однако, необходимо разработать эффективные методы обработки и анализа данных, а также обеспечить защиту персональных и конфиденциальных данных, чтобы использование больших данных в науке было этичным и надежным.

Социальные исследования и государственное управление. Большие Данные используются для анализа данных о населении, транспортных потоках, экологической обстановке и других факторах, что позволяет принимать обоснованные решения в области государственного управления и планирования развития. Применение Больших Данных в социальных исследованиях и государственном управлении имеет огромное значение для понимания общественных явлений, принятия информированных решений и разработки эффективных стратегий в различных областях [12]. Вот некоторые примеры применения Больших Данных в этой области.

1) *Анализ социальных трендов и общественного мнения*. Большие Данные позволяют анализировать социальные медиа, новостные и другие публичные источники информации для выявления социальных трендов, общественного мнения и настроений. Это помогает исследователям и правительственным органам понять общественные предпочтения, потребности и проблемы, а также принимать меры для улучшения политик и программ. Большие Данные предоставляют возможность анализировать социальные сети, медиасообщества, электронные документы и другие источники информации для изучения социальных явлений, поведения людей, мнений и тенденций. Это помогает социологам, психологам, экономистам и другим исследователям понять социальные процессы, развитие общества, прогнозировать социальные изменения и разрабатывать политики на основе фактических данных.

2) *Прогнозирование и предотвращение социальных кризисов*. С помощью анализа Больших Данных можно прогнозировать социальные кризисы, такие как нарушения общественного порядка, конфликты и экономические кризисы. Анализ данных о прошлых событиях, социальных индикаторах и других факторах позволяет предупреждать и предотвращать потенциальные проблемы, а также разрабатывать стратегии по управлению кризисными ситуациями.

3) *Улучшение государственных услуг и решение общественных проблем*. Большие Данные используются для анализа эффективности государственных программ и услуг, выявления узких мест и определения областей, требующих улучшения. С помощью анализа данных можно определить оптимальные стратегии по распределению ресурсов, улучшению доступности услуг и решению общественных проблем, таких как бедность, безработица и неравенство.

4) *Принятие решений, на основе данных*. Большие Данные помогают правительственным органам принимать информированные решения в различных областях, таких как экономика, здравоохранение, образование и транспорт. Анализ данных позволяет оценить эффективность политик, предсказать результаты различных сценариев и разрабатывать стратегии для достижения поставленных целей.

5) Развитие умных городов и цифровых государств. Большие Данные играют ключевую роль в развитии умных городов и цифровых государств. Анализ данных о жизни горожан, экологических показателях, инфраструктуре и других аспектах позволяет оптимизировать управление городскими ресурсами, повысить уровень жизни граждан, обеспечить устойчивое развитие и повысить эффективность государственных услуг.

Применение Больших Данных в социальных исследованиях и государственном управлении помогает принимать более обоснованные решения, повышать эффективность и качество государственных программ и услуг, а также улучшать качество жизни граждан. Однако, необходимо учитывать этические и конфиденциальные аспекты использования данных, чтобы обеспечить защиту личной информации и поддерживать доверие общества к использованию данных в целях общественного блага.

Финансы. Большие Данные используются для анализа финансовых транзакций, предсказания трендов на рынке и обнаружения мошенничества. Это позволяет банкам принимать более точные решения и снижать риски. Банки, работая с транзакционной информацией, используют распределенные вычисления, что полезно для выявления мошенничества и улучшения работы сервисов [13].

Производство. Большие Данные используются для оптимизации производственных процессов, предотвращения отказов оборудования, прогнозирования спроса и повышения качества продукции. Анализ Больших Данных также может помочь компаниям сократить издержки на производство [2, 14].

Телекоммуникации. Большие Данные позволяют проводить анализ поведения клиентов, оптимизировать сетевую инфраструктуру, предсказывать спрос и улучшать качество обслуживания. Задачи, которые решаются с использованием Больших Данных в телекоммуникациях

– разработка и запуск новых продуктов/услуг;

– формирование предикативных предложений;

– прогнозирование жизненного цикла продуктов/услуг;

– ускорение обслуживания клиентов и принятия решений о предоставлении/отказе в услуге.

– снижение затрат на поддержку.

Маркетинг. Большие Данные используются для анализа поведения потребителей, чтобы определить, какие продукты и услуги будут наиболее востребованы в будущем [15]. Также, Большие Данные используются для измерения эффективности рекламных кампаний и принятия решений по оптимизации маркетинговых стратегий. Большие данные обеспечивают более глубокий и объективный анализ рынка и конкурентной среды. Компании могут анализировать данные из социальных медиа, интернет-форумов, отзывов потребителей и других источников, чтобы получить информацию о мнениях и предпочтениях потребителей, а также о реакции на маркетинговые акции и рекламные кампании.

Применение Больших Данных в этих отраслях позволяет выявлять скрытые закономерности, прогнозировать события, принимать обоснованные решения и достигать конкурентных преимуществ. Это лишь часть сфер, где растет востребованность аналитики Больших Данных. По данным Data Age Report человечество сформировало порядка 51 зеттабайта информации в 2020 году (рис. 4). К 2025 году объем этих данных вырастет до 175 зеттабайтов ежегодно [16].

По оценкам, к 2025 году во всем мире будет создаваться 463 экзабайта данных каждый день, что эквивалентно 212 765 957 DVD в день [17].

Рис. 4. Динамика роста объема данных в мире

Такие объемы невозможно обрабатывать традиционными базами данных, поэтому компании видят решение перехода на Большие Данные не просто как обработку огромных объемов, но и как повышение конкурентоспособности, увеличения лояльности покупателя к своему продукту и привлечения новых. Согласно институту статистических исследований и экономики знаний НИУ ВШЭ в 2021 г. технологии сбора, обработки и анализа больших данных в России применяли 25,8% организаций – на 3,4% больше, чем в 2020 г. Самым популярным источником оказались веб-сайты компаний: их данные собирают 9,2% компаний. Немного менее востребованы данные учетных систем организации (ERP, CRM и др.) и социальных сетей (8 и 7,2% соответственно). Помимо этого, анализируются данные операторов сотовой связи (6,7%), информация с цифровых датчиков и радиочастотных меток (6,3%).

Главный поставщик Больших Данных сегодня – это государство, и именно оно в дальнейшем будет главным потребителем их анализа. Госструктуры анализируют Большие Данные для повышения безопасности граждан и совершенствования городской инфраструктуры, улучшения работы сфер ЖКХ и общественного транспорта. Государству Большие Данные от сотовых операторов помогают прогнозировать перемещения людей и планировать, например, развитие транспортной сети, которое будет лучше удовлетворять потребностям людей. В частности, московские власти еще с 2019 г. применяют их при проектировании и строительстве транспортной инфраструктуры, для регулирования пассажирских тарифов и обеспечения безопасности дорожного движения. Компания «С-Плэтформс» планирует запустить в особой экономической зоне Москвы специализированный центр обработки Больших Данных. Площадь нового дата-центра составит около 850 м^2, мощность – ориентировочно 2,5 мегаватта с размещением 132 стоек. Здесь планируется создать 27 рабочих мест [18].

Спрос на Большие Данные продолжает расти по всему миру, поскольку все больше компаний и организаций осознают важность анализа данных для своего бизнеса. Ниже приведены некоторые тренды, связанные с развитием спроса на Большие Данные [19]:

– *рост объемов данных*. Объемы данных продолжают расти по всему миру, и компании нуждаются в инструментах для обработки и анализа этих данных. Согласно прогнозам, объемы данных будут расти в 3-4 раза каждые два года;

– *развитие технологий обработки Больших Данных*. Развитие технологий обработки Больших Данных позволяет компаниям анализировать данные быстрее и более эффективно. В частности, машинное обучение [20-23] и искусственный интеллект [24] позволяют компаниям получать более точные результаты анализа данных;

– *рост спроса на облачные сервисы*. Облачные сервисы предоставляют компаниям доступ к инфраструктуре для обработки Больших Данных без необходимости владения собственными серверами или вычислительными мощностями. Это позволяет компаниям снизить издержки на обработку данных и ускорить процесс анализа данных;

– *развитие Интернета вещей* (IoT). Развитие IoT приводит к возникновению большого количества данных, которые могут быть использованы для повышения эффективности бизнеса [25]. Компании используют данные, полученные от IoT-устройств, для оптимизации производства, управления запасами и повышения качества продукции;

– *искусственный интеллект*. Совмещение Больших Данных с технологиями искусственного интеллекта, такими как машинное обучение и глубокое обучение, позволяет создавать более интеллектуальные и автономные системы, способные анализировать и принимать решения на основе больших объемов данных.

– *развитие визуализации данных*. Визуализация Больших Данных становится все более важным инструментом для исследования и визуального представления сложных и объемных данных. Продвижение визуализации данных и разработка новых методов визуализации помогают улучшить восприятие и понимание информации, заключенной в больших объемах данных;

– *рост спроса на аналитиков данных*. В связи с ростом объемов данных и развитием технологий обработки Больших Данных увеличивается спрос на специалистов по анализу данных. Компании нуждаются в аналитиках данных, которые могут анализировать большие объемы данных и предоставлять бизнес-аналитику, необходимую для принятия решений.

Проблемы внедрения Больших Данных в рамках третьей платформы информатизации

Одной из важных проблем при работе с Big Data является *классификация* [1]. Несмотря на то, что есть способы классифицировать данные и применяемые к ним технологии, свойственный Большим Данным плюрализм не позволяет создать единые направления и методы для работы с ними т. к. Большие Данные могут иметь различные форматы и структуры. Они включают структурированные данные (например, данные в таблицах и базах данных), полуструктурированные данные (например, данные в формате XML или JSON) и неструктурированные данные (например, текстовые документы, изображения, аудио и видеофайлы).

Проблемы, связанные с классификацией Больших Данных

Рассмотрим имеющиеся способы классификации общего понятия Больших Данных без привязки к какому-то конкретному набору. Большие Данные создают характерные особенности, которые не разделяются традиционными наборами данных. Эти особенности можно свести к трем основным группам: объем, скорость обработки, многообразие – Volume, Velocity и Variety.

– *Volume* (*объём*) – представляет собой величину физического объёма данных. К 2020 г. общий объем информации, созданный в цифровой среде, достиг 51 зеттабайтов [16]. В России он тоже будет расти в любом случае, хоть и не такими высокими темпами, какими мог бы без ухода иностранных игроков, считает гендиректор Института исследований интернета Карен Казарян: «"Яндекс" и VK так или иначе продолжат работать и развиваться». В то же время он полагает, что опасения людей, связанные с введением цифровых повесток, едва ли скажутся на объеме рынка, поскольку «вопрос не в сборе данных, а в их использовании» [26]. В 2022 году практически половину интернет-трафика в глобальном масштабе сгенерировали боты [27]. Об этом говорится в исследовании компании Imperva, результаты которого были обнародованы 10 мая 2023-го. По оценкам, в 2022 г. на различные автоматизированные системы пришлось 47,4% от общего объёма данных, переданных в Сети. Это на 5,1% больше по сравнению с предыдущим годом. Вместе с тем доля трафика, сгенерированного людьми, сократилась до 52,6% – это самый низкий показатель за восемь лет (к началу 2023 г.).

В отчёте Imperva говорится о продолжающемся росте потока данных от так называемых «плохих» ботов, то есть, автоматизированных инструментов, которые могут применяться для рассылки спама, проведения DDoS-атак и организации других злонамеренных кампаний. В 2022 году доля трафика от таких систем составила 30,2%, что на 2,5% больше по сравнению с 2021 г. В то же время боты, не представляющие угрозы для пользователей, нарастили долю веб-трафика до 17,3% – плюс 2,7% в годовом исчислении. Количество атак с захватом учётных записей в течение 2022 года увеличилось на 155%, чему способствовали утечки персональной информации. Приблизительно 17% всех атак на программные интерфейсы приложений (API) были осуществлены «плохими» ботами. Чаще всего атаки со стороны автоматизированных систем осуществлялись на ресурсы в сфере туризма (24,7%), розничной торговли (21%) и финансовых услуг (12,7%). Обычные инструменты хранения и анализа не способны справляться с таким объемом данных [1-6].

– *Velocity* (*скорость*) – подразумевает под собой как скорость прироста информации, так и необходимость высокоскоростной обработки и получения результатов. Указанные выше объемы данных поступают в обработку в режиме реального времени, в отличие от традиционной обработки пакета данных. Это означает, что они накапливаются моментально, при этом не имеет значения продолжительность потока самих данных. Таким образом, Big Data не только фиксирует потоки данных, но и производит их запись и обработку в таком виде, чтобы не было потерь. Примером потоковой обработки данных является сервис YouTube, проводящий анализ данных пользователей, исходя не только из просмотренных полностью видеозаписей и трансляций, но из пропущенных пользователями материалов и воспринятых ими в качестве ненужных. Для целей авторов каналов YouTube дополнительно предлагает услуги по сбору данных об интересах зрителей, географических особенностях, контентных предпочтениях, предложения по целевой аудитории.

– *Variety* (*многообразие*) – возможность одинаковой и одновременной обработки различных типов данных: структурированных и полуструктурированных, неструктрированных. Big Data формируется из различных источников и в виде множества разнообразных форматов данных (видеоданные, фотографии, звуковые записи, текстовые сообщения, файлы транзакций, комментарии, использование ссылок и фиксация просмотров страниц и т.д.). Таким образом, термин Big Data не относится исключительно к «большим данным» в понимании объема. Он значительно шире, поскольку включает в себя также большие скорости поступления данных и большое разнообразие источников и форматов получаемой информации.

В некоторых случаях используют дополнительные признаки Big Data [1-6].

– *Veracity* (*достоверность*) – представляет собой набор истинной информации, учёт которой при обработке массивов данных является наиболее важным. IDC интерпретирует «Veracity» как value с точки зрения важности экономической целесообразности обработки соответствующих объёмов в соответствующих условиях, что отражено также и в определении Больших Данных от IDC. Из-за большого объема и вариативности источников поступающих данных сложно проконтролировать достоверность Big Data. Соответствие, точность и правдивость получаемой информации могут быть подтверждены только в результате тщательного анализа и сопоставления.

– *Variability* (*переменчивость*) – способность данных терять свою актуальность со временем. При обработке и сопоставлении исходного значения полученных данных может меняться, то есть зависит от определенного контекста. В первую очередь данный признак проявляется при работе с речевыми и текстовыми данными. Необходимо определить смысловую нагрузку исходя не только из прямого значения, но и из контекста. Может является частным показателем для жизнеспособности, если рассматривать их как единое целое и в одной системе классификации.

– *Visualization* (*визуализация*) – характеризует набор данных в зависимости от степени удобства их представления и графической интерпретации. Полученные в результате сбора данные непригодны для восприятия человеком. Поэтому требуется их обработка в доступной форме – процедура визуализации. Характерным примером визуализации данных является построение графиков и диаграмм, отображающих результаты анализа данных. Важным является возможность самостоятельной настройки визуализации Больших Данных, в зависимости от поставленных целей и задач.

– *Value* (*ценность*) – показатель, характеризующий важность и необходимость выборки данных при работе с ними над решением конкретных практических задач и/или *Viability* (*жизнеспособность*) – характеризует данные в зависимости от времени их актуальности. На ценность влияют указанные выше признаки Больших Данных: тщательный и точный анализ данных, актуальность информации и полученные в результате визуализации выводы.

Далее кратко укажем основные группы проблем, связанные с хранением, выделением знаний из больших данных и вычислительной сложностью обработки, масштабируемостью и визуализацией больших данных, а также информационной безопасностью.

Проблемы, связанные с хранением Больших Данных

Первая группа проблем хранения Больших Данных (особенно когда данные представлены в разных форматах) в первую очередь связана с их объемом, требующим дорогостоящих технологий для хранения и обработки [28, 29]. Чтобы объединить данные и эффективно их обрабатывать, требуется не только работа по приведению их в пригодный для работы вид, но и определенные аналитические инструменты (системы). Проблема хранения возникает также и в силу увеличения скорости нарастания объемов новых данных в последние годы [16, 17, 26, 27]. В силу недостатка места для их хранения они либо удаляются, либо не записываются вовсе. В связи с этим, возрастает роль носителей информации и скорости её записи и чтения для доступности Больших Данных с целью их анализа. Несмотря на достижения в этой области, такие как, например, распространение твердотельных накопителей, необходимая производительность накопителей для обработки больших данных до сих пор не достигнута. В мае 2017 года Министерство обороны Российской Федерации объявило о создании сверхстабильной оптической памяти, технология которой основана на создании нанорешеток с определенными свойствами путем облучения кристаллов кварца лазерным излучением с заданными параметрами. Сообщается, что новая разработка обладает следующими характеристиками [30].

– долговечность – сохранение физико-химических свойств и способность хранить записанные данные при комнатной температуре в течение неограниченно длительного срока – сотен тысяч лет и более;

– устойчивость к высоким температурам – не менее 800ºС;

– высокая радиационная и химическая стойкость;

– повышенная емкость и сопоставимые с современными носителями скорости записи (от 10 Мбит/с) и считывания (от 100 Мбит/с) данных.

Кроме объёмов производимой информации, в последние годы также нарастает и их *разнообразие* [16, 17, 26, 27], что значительно усложняет задачи классификации, упорядочивания и анализа Больших Данных. Последние технологии, такие как Hadoop [31, 32] и MapReduce [33, 34], позволяют собирать большие объёмы полуупорядоченных и неупорядоченных данных за приемлемое время. Hadoop основан на модели MapReduce и позволяет работать с данными объемом до нескольких петабайт. Hadoop состоит из двух основных компонентов: Hadoop Distributed File System (HDFS) и MapReduce. HDFS – это распределенная файловая система, которая позволяет хранить большие объемы данных на нескольких узлах компьютерной сети. MapReduce – это модель программирования, которая использует параллельные вычисления для обработки больших объемов данных. Особенностями Hadoop, которая, по сути, является не обычной классической базой данных, а файловой системой, организованной в так называемое «озеро данных», где хранятся данные из различных источников. При этом информация в таком озере физически распределена по кластеру серверов и доступна через

различные интерфейсы (API) или прикладные слои, каждый из которых необходимо защищать.

Проблемы выделения знаний из Больших Данных и вычислительной сложности их обработки

Вторая группа проблем – выделения знаний из Больших Данных и вычислительная сложность их обработки. Огромные массивы не всегда возможно хранить на одном сервере, что в свою очередь приводит к применению технологий распределённых систем – совокупность взаимосвязанных автономных компьютеров и их вычислительных мощностей. Однако, распределённое хранение также вызывает сложности при выборки данных и составлении алгоритмов их обработки. Главной проблемой при анализе является устранение несоответствий и неопределённости, которые присутствуют в наборах данных [35, 36]. Ее усугубляет сложность структуризации, сортировки, распределения при составлении выборок и поиске конкретного элемента из общей системы [1].

Несмотря на то, что попытки преодоления вычислительной сложности реализуются в большинстве случаев обработки наборов Больших Данных, единого метода, применимого ко всем случаям, до сих пор не существует. Имеющиеся инструменты анализа имеют слишком низкую производительность и не в состоянии эффективно справляться с несоответствиями, неопределённостью и вычислительной сложностью, которые возникают при обработке наборов Больших Данных. Разработки в этой области ведутся с использованием машинного обучения [20-23].

Выделение и представление знаний из Больших Данных – главная задача их обработки. Она включает в себя несколько подзадач, таких как архивирование, управление, сохранение, поиск и представление знаний. Отсутствие эффективных алгоритмов обработки, учитывающих объём хранилища данных, структуру и методы поиска необходимого элемента (ячейки памяти) также является одной из проблем этого класса. Алгоритмы, которые используются для решения этих задач, основаны, по большей части, на теории нечётких множеств и нечёткой логики, которые в настоящее время активно развиваются.

Подбор данных для обработки и алгоритм анализа может стать не меньшей проблемой, так как отсутствует понимание, какие данные следует собирать и хранить, а какие можно игнорировать. Здесь появляется проблема шумов и их учёт при работе с датасетами [1]. Дело в том, что в структурированных наборах данных, представляющих собой реляционные SQL системы, отклонения от общей структуры (по форме данных, их содержанию) считаются выбросами, и зачастую не учитываются (отбрасываются) при составлении общих выборок. Однако в случае с Больших Данных выбросы и отклонения зачастую содержат в себе наиболее важную информацию, а сам большой объём данных формируется с целью выявить эти самые отклонения. Небольшая по размеру их выборка (по сравнению с большим объёмом общего хранилища) имеет наибольшую ценность в практической и исследовательской деятельности. И очень важно при обработке датасета их не пропустить и не отбросить. Это первостепенное противоречие, связанное с проблемой наличия шумов.

Проблемы масштабируемости и визуализации Больших Данных

Третья группа проблем – масштабируемость и визуализация Больших Данных. В последнее время исследования в области Больших Данных позволили добиться ускорения их обработки, на фоне увеличения производительности процессоров по закону Мура. Несмотря на это, объёмы Больших Данных растут гораздо быстрее, чем производительность процессоров.

В связи с этим, ставится задача распараллеливания вычислений между разными процессорами, в том числе между различными ядрами одного и того же процессора. Методы и алгоритмы параллельных вычислений являются одной из областей исследования.

Цель визуализации набора Больших Данных – дать аналитикам адекватное представление об их свойствах, помочь правильно их интерпретировать. Визуализация позволяет превратить большой массив данных в графики или изображения, которые дадут аналитикам интуитивное представление об их содержании. Современные инструменты визуализации обладают неудовлетворительной производительностью, функционалом и временем отклика, что также является проблемой исследований.

Проблемы информационной безопасности Больших Данных

Четвёртая группа проблем – проблемы информационной безопасности, связанные с Большими Данными. Информационная безопасность становится проблемой при анализе Больших Данных. В любом случае при передаче, обработке и хранении данных должны быть обеспечены основные услуги безопасности: целостности, доступности и конфиденциальности с учетом требований Политики информационной безопасности организации и регулятора [37-40]. Перечислим основные проблемы информационной безопасности Больших Данных [41].

Отсутствие практики по работе с Большими Данными и её защите – это новая парадигма хранения и обработки данных. Администраторы службы информационной безопасности не всегда понимают, что именно происходит внутри кластера с Большими Данными, каковы угрозы и уязвимости новых технологий. Методологии по защите информационных систем классической двух-, трёхзвенной архитектур оказываются не применимы к новым технологиям.

Отсутствие методологий по защите Больших Данных. Различные организации публикуют свои методологии и рекомендации, однако уровня ISO пока ни одна из них не достигла.

Отсутствие стандартов по защите Больших Данных. На сегодняшний момент над созданием стандартов по защите Big Data работает несколько международных рабочих групп (WG9 под эгидой комитета ISO JTC 1, Big Data Working group от сообщества Cloud Security Alliance, Big Data – NIST SP1500-4: Big Data Security and Privacy), однако, ни одного опубликованного стандарта на сегодня нет.

Большая экосистема Больших Данных. Экосистема Больших Данных чересчур активно развивается и слишком быстро растёт, что усложняет её стандартизацию.

Указанные проблемы требуют индивидуального решения безопасности Больших Данных. Например, в СБЕРе безопасность Больших Данных реализована на двух уровнях: идентификация и классификация информации (объектам защиты устанавливаются метки конфиденциальности данных в озере данных); обеспечение безопасности (применение мер обеспечения безопасности к объектам защиты). Например, административные, физические и технические меры обеспечения защиты. Требования к мерам защиты можно найти в различных сборниках стандартов, например, в ISO 27001. В настоящее время ведутся работы по интеграции классических систем безопасности и полномасштабных аналитик на базе «больших данных» для решения задач обеспечения информационной безопасности [42] и др.

Еще одна проблема безопасности Больших Данных носит этический характер. А именно: чем сбор данных (особенно без ведома пользователя) отличается от нарушения границ частной жизни? Так, информация, сохраняемая в поисковых системах Google и Яндекс, позволяет им постоянно дорабатывать свои сервисы, делать их удобными для пользователей и создавать новые интерактивные программы. Поисковики записывают каждый клик пользователя в Интернете, им известен его IP-адрес, геолокация, интересы, онлайн-покупки, личные данные, почтовые сообщения и прочее, что, к примеру, позволяет демонстрировать контекстную рекламу в соответствии с поведением пользователя в Интернете. При этом согласия на это не спрашивается, а возможности выбора, какие сведения о себе предоставлять, не дается. То есть по умолчанию в Большие Данные собирается все, что затем будет храниться на серверах данных сайтов.

Заключение

Технология Больших Данных является неотъемлемой частью современного информационного общества. Она предоставляет большие возможности для обработки, анализа и извлечения знаний из огромных объемов данных. С появлением новых технологий и инструментов, таких как машинное обучение, искусственный интеллект, блокчейн и интернет вещей, возникают новые возможности для обработки и анализа Больших Данных в сферах здравоохранения и медицины, финансов, научных и социальных исследований, транспорта и логистики, государственного управления, производства, телекоммуникаций, маркетинга и др.

Третья платформа информатизации играет важную роль в работе с Большими Данными, предоставляя инфраструктуру и инструменты для эффективного управления и анализа данных. Однако с появлением Больших Данных возникают и ряд проблем, связанных с их обработкой, хранением, анализом и защитой.

Заглядывая вперед, можно сказать, что исследования в области технологий Больших Данных будут развиваться в направлении разработки новых технологий для их хранения, эффективных алгоритмов обработки и выделения знаний, защиты и др. на базе искусственного интеллекта и машинного обучения, а также аналитики Больших Данных.

Литература

1. *Менщиков А.А., Перфильев В.Э., Федосенко М.Ю., Фабзиев И.Р.* Основные проблемы использования больших данных в современных информационных системах // Столыпинский вестник. 2022. №1. С. 316-329.

2. *Федорова Л.А., Ху Гуйюй, Хуан Сяоянь, Землякова С.А.* Применение технологий Big Data в деятельности современных предприятий // Вестник Алтайской академии экономики и права. 2020. № 9-2. С. 322-329.

3. *Клеменков П.А., Кузнецов С.Д.* Большие данные: современные подходы к хранению и обработке // Труды Института системного программирования РАН. 2012. № 4. С. 143-155.

4. *Тесленко И.Б., Губернаторов А.М., Дигилина О.Б., Крылов В.Н.* Big Data = Большие данные. Владимир: Владим. гос. ун-т им. А. Г. и Н. Г. Столетовых. 2021, 121 с.

5. Big Data // Gartner Glossary [Электронный ресурс]. URL: https://www.gartner.com/en/information-technology/glossary/big-data (дата обращения 17.05.2023).

6. *Абдыкаримова А.Т.* Big Data: проблемы и технологии // Международный журнал гуманитарных и естественных наук. 2019. № 5-1. С. 55-57. URL: https://cyberleninka.ru/article/n/big-data-problemy-i-tehnologii (дата обращения: 17.05.2023).

7. *Колбанёв М.О., Татарникова Т.М., Верзун Н.А.* Технологическая платформа четвертой промышленной революции // Геополитика и безопасность. 2016. № 2 (34). С. 73-77.

8. *Найдич А.* «Третья платформа» – платформа трансформации ИТ. // КомпьютерПресс. 2013. URL: https://compress.ru/article.aspx?id=24166 (дата обращения: 10.09.2023).

9. *Гуревич М.Н.* Большие данные: исследование и анализ // Информатика и вычислительная техника. 2017. № 3. С. 48-57.

10. *Петросян М.К., Михнев И.П., Новикова А.А.* Большие данные (big data) и новые технологии будущего для обработки глобальной информации // II Международная научно-практическая конференция «Научные исследования и современное образование». 2018. С. 1-8.

11. *Авраменко А.А.* Анализ данных в условиях больших данных // Вестник Пермского университета. Серия: Математика. Механика. Информатика. 2017. № 13 (4). С. 12-20.

12. *Романенко Е. В.* Место big data в современной социально-экономической жизни общества // Инновационная наука. 2016. № 4. С. 143-145.

13. Банки, ретейл, медицина: кто использует Data Mining и для чего [Электронный ресурс]. СПб., 2021. URL: https://trends.rbc.ru/trends/industry/61b359739a7947c7376ef7ce. (Дата обращения: 10.07.2023).

14. *Антонов А.А., Потапов А.В.* Реализация технологий больших данных для анализа данных в управлении предприятием // Вестник Южно-Уральского государственного университета. Серия: Экономика и менеджмент. 2018. № 12 (2), С. 81-88.

15. *Сергеев В.И.* Большие данные в маркетинге: анализ и применение // Маркетинг в России и за рубежом. 2018. № 2. С. 32-44.

16. Большие данные (Big Data) мировой рынок. [Электронный ресурс]. URL: https://www.tadviser.ru/index.php/ (дата обращения: 10.07.2023).

17. *Desjardins J.* How much data is generated each day? // World Economic Forum Articles. 2019. [Электронный ресурс]. URL: https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/ (дата обращения: 09.07.2023).

18. Дата-центр для работы с большими данными построят в Москве [Электронный ресурс]. URL: https://www.iksmedia.ru/news/5890168-Datacentr-dlya-raboty-s-bolshimi.html (дата обращения: 09.07.2023).

19. *Николаева Н.В.* Большие данные: применение в науке и бизнесе // Информационные технологии в образовании, науке и промышленности. 2017. № 2 (16). С. 81-88.

20. *Жернакова Е.В., Николаев А.А.* Анализ больших данных с использованием технологии машинного обучения // Информационные технологии. 2018. № 4. С. 21-28.

21. *Дьяконов А.Г.* Большие данные и машинное обучение // Автоматика и телемеханика. 2018. № 79 (10). С. 1563-1575.

22. *Камалбекова А.А., Стрекаловский А.О.* Анализ больших данных с помощью методов машинного обучения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. № 44 (4). С. 48-60.

23. *Патрин Е.С., Калинова О.Ю.* Большие данные: статистический анализ и машинное обучение в экономике // Экономика региона. 2019. № 15 (3). С. 736-748.

24. *Полетаев А.И., Азаров В.В.* Большие данные и искусственный интеллект в экономике // Вестник Московского государственного университета им. М. В. Ломоносова. Серия: Экономика. 2019. № 5. С. 39-52.

25. *Тимошин И.А.* Большие данные и интернет вещей: анализ и применение // Вестник Пензенского государственного университета. 2019. № 4. С. 146-154.

26. Большие данные прошли переоценку // Коммерсантъ [Электронный ресурс]. URL: https://www.kommersant.ru/doc/5939856 (дата обращения: 09.07.2023).

27. Интернет-трафик (мировой рынок) // TAdviser [Электронный ресурс]. URL: https://www.tadviser.ru/index.php/ (дата обращения: 09.07.2023).

28. *Казаковцева О.А., Черкасова Т.В.* Большие данные: проблемы обработки и хранения информации // Известия Уральского государственного университета. 2019. № 187 (3). С. 68-81.

29. *Лукьянов В.И., Рябцев В.И.* Большие данные: проблемы анализа и хранения информации. // Вестник Удмуртского университета. Серия: Математика. Механика. Компьютерные науки. 2019. № 29 (2). С 220-232.

30. Технология сверхстабильной оптической памяти разработана в интересах Минобороны России. [Электронный ресурс]. URL: https://function.mil.ru/news_page/country/more.htm?id=12122289@egNews (дата обращения: 10.07.2023).

31. *Гусев К.Ю., Кораблина Т.М., Степанова Е.Н., Жижелева Е.А.* Применение технологии Hadoop для обработки больших данных // Интернет-математика. 2015. Т. 11. № 3. С. 131-141.

32. *Астапенко Т.С., Соколин Д.Д.* Проблемы безопасности системы обработки больших данных Hadoop // Решетневские чтения. 2018. С. 314-315 [Электронный ресурс]. URL: https://cyberleninka.ru/article/n/problemy-bezopasnosti-sistemy-obrabotki-bolshih-dannyh-hadoop (дата обращения: 10.07.2023).

33. *Давыдова О.В., Савина И.А., Белоусова О.И., Петров А.В., Гарифуллин Р.Р.* Использование технологии MapReduce для обработки больших данных // Автоматика и телемеханика. 2015. № 11. С. 128-135.

34. *Некратюк А.А., Сафарьян О.А.* Использование метода MapReduce в big data // Молодой исследователь Дона. 2020. №3 (24). [Электронный ресурс]. URL: https://cyberleninka.ru/article/n/ispolzovanie-metoda-mapreduce-v-big-data (дата обращения: 10.07.2023).

35. *Кукарцева О.И., Корнеева Е.А., Храмков В.В.* The importance of big data processing. // Актуальные проблемы авиации и космонавтики. 2022. Т. 2. С. 208-209.

36. *Колтаков В.П.* Большие данные и аналитика в информационно-аналитических системах // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. 2019. № 2. С. 9-17.

37. *Мошак Н.Н.* Безопасность информационных систем: учебно-методическое пособие. СПб.: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2020. 73 с.

38. *Мошак Н.Н.* Защита информационных систем: учебно-методическое пособие. СПб.: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2020. 154 с.

39. *Мошак Н.Н., Птицына Л.К.* Защищенные информационные системы: учебно-методическое пособие. СПб: Санкт-Петербургский государственный университет Телекоммуникаций им. профессора М. А. Бонч-Бруевича, 2020. 216 с.

40. *Мошак Н.Н.* Основы управления информационной безопасностью: учебно-методическое пособие. СПб.: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2022. 141 с. ISBN 978-5-8088-1711-1.

41. Защита Big Data: проблемы и решения. [Электронный ресурс]. URL: https://www.itworld.ru/cionews/security/158148.html (дата обращения: 10.07.2023).

42. *Мурадова Г.* Вопросы информационной безопасности «больших данных». 2015. С. 228-231. [Электронный ресурс]. URL: https://ict.az/uploads/konfrans/2_konfrans/63.pdf (дата обращения: 11.07.2023).

THIRD PLATFORM FOR INFORMATIZATION AND BIG DATA

NIKOLAI N. MOSHAK
St. Petersburg, Russia

SABINA R. RUDINSKAYA
St. Petersburg, Russia

ALEXEY A. GRUZDEV
St. Petersburg, Russia

KEYWORDS: *information analysis, Big Data, virtualization, clustering, data processing, third informatization platform.*

ABSTRACT

Introduction: despite the fact that the introduction of Big Data in many areas of life is an inevitable reality of the near future, there are difficulties with the classification of information, the allocation of knowledge, storage, security and scalability issues. The paper describes the concept of the third platform of informatization and Big Data, which play an important role in this process, is described. The problems associated with the need for efficient processing and analysis of Big Data are posed. The spheres of life in which new technologies are being actively introduced and developed are shown and considered. **The purpose of the study** is to discuss the trends in the implementation of Big Data within the framework of the third informatization platform. The problems of Big Data and some ways to solve them are investigated. The interrelations and patterns in the field of Big Data and their impact on the development of information technologies are considered. **Research methodology:** analytical research methods based on the analysis of data and concepts of the third informatization platform are used. **The results of the study:** a deep analysis was carried out for full immersion in the issue of the third platform of informatization and the concept of Big Data. The advantages of introducing technology components into the spheres of life are shown, as well as the problems of Big Data and some ways of their solution are studied. The main theoretical results of the study are presented, confirming the interrelationships and patterns between the elements of developing information technologies. The practical significance of the third informatization platform and its components is discussed. Recommendations related to the application of these technologies in real conditions are proposed, as well as their mutual influence of technologies to achieve more effective results.

REFERENCES

1. A.A. Menshchikov, V.E. Perfil'ev, M.U. Fedosenko, I.R. Fabziev, "The main problems of use of big data in modern information systems," *Stolypinskii vestnik* [Stolypin Bulletin]. 2022. No. 1, pp. 316-329. (In Rus)

2. L.A. Fedorova, KHu Guiiui, KHuan Siaoian', S.A. Zemliakova, "Application of big data technologies in the activities of modern enterprises," *Vestnik Altaiskoi akademii ekonomiki i prava* [Bulletin of the Altai Academy of Economics and Law]. 2020. No. 9-2, pp. 322-329. (In Rus)

3. P.A. Klemenkov, S.D. Kuznetsov, "Big data: modern approaches to storage and processing," *Trudy Instituta sistemnogo programmirovaniia RAN* [Proceedings of the Institute of System Programming of the Russian Academy of Sciences]. 2012. No. 4, pp. 143-155. (In Rus)

4. I.B. Teslenko, A.M. Gubernatorov, O.B. Digilina, V.N. Krylov, "Big Data = Big Data," Vladimir. Vladimir State University named after A. G. and N. G. Stoletov Publ. 2021. 121 p. (In Rus)

5. Big Data. Gartner Glossary. Avaliable at: https://www.gartner.com/en/information-technology/glossary/big-data (date of access 17.05.2023).

6. A.T. Abdykarimova, "Big data: problems and technologies," *Mezhdunarodnyi zhurnal gumanitarnykh i estestvennykh nauk* [International Journal of Humanities and Natural Sciences]. 2019. No. 5-1, pp. 55-57. Avaliable at: https://cyberleninka.ru/article/n/big-data-problemy-i-tehnologii (date of access 17.05.2023). (In Rus)

7. M.O. Kolbanev, T.M. Tatarnikova, N.A. Verzun, "The technological platform of the fourth industrial revolution," *Geopolitika i bezopasnost'* [Geopolitics and security]. 2016. No. 2 (34), pp. 73-77 (In Rus)

8. A. Naidich, "The Third Platform is an IT transformation platform," Komp'iuterPress [ComputerPress]. 2013. Avaliable at: https://compress.ru/ article.aspx?id=24166 (date of access 10.09.2023). (In Rus)

9. M.N. Gurevich, "Big data: research and analysis," *Informatika i vychislitel'naia tekhnika* [Computer science and engineering]. 2017. no. 3, pp. 48-57 (In Rus)

10. M.K. Petrosian, I.P. Mikhnev, A.A. Novikova, "Big data and new technologies of the future for processing global information," *Proceedings of the 2th International Scientific Conference "Nauchnye issledovaniia i sovremennoe obrazovanie"* [II International Scientific and Practical Conference "Scientific research and modern Education]. 2018, pp. 1-8. (In Rus)

11. A.A. Avramenko, "Data analysis in a big data environment," *Vestnik Permskogo universiteta. Seriia: Matematika. Mekhanika. Informatika* [Bulletin of Perm University. Series: Mathematics. Mechanics. Computer science]. 2017. No. 13 (4), pp. 12-20. (In Rus)

12. E.V. Romanenko, "The place of big data in the modern socio-economic life of society," *Innovatsionnaia nauka* [Innovative science]. 2016. No. 4, pp. 143-145. (In Rus)

13. Banks, retail, medicine: who uses Data Mining and for what [Online resource]. Saint Petersburg, 2021. Avaliable at: https://trends.rbc.ru/trends/industry/61b359739a7947c7376ef7ce. (date of access 10.07.2023). (In Rus)

14. A.A Antonov, A.V. Potapov, "Implementation of big data technologies for data analysis in enterprise management," *Vestnik IUzhno-Ural'skogo gosudarstvennogo universiteta. Seriia: Ekonomika i menedzhment* [Bulletin of the South Ural State University. Series: Economics and Management]. 2018. No. 12 (2), pp. 81-88. (In Rus)

15. V.I. Sergeev, "Big Data in marketing: analysis and application," *Marketing v Rossii i za rubezhom* [Marketing in Russia and abroad]. 2018. No. 2, pp. 32-44. (In Rus)

16. Big Data (Big Data) global market [Online resource]. Avaliable at: https://www.tadviser.ru/index.php/ (date of access 10.07.2023). (In Rus)

17. J. Desjardins, "How much data is generated each day? World Economic Forum Articles. 2019," [Online resource]. Avaliable at: https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/ (accessed: 09.07.2023).

18. Data-tsentr dlia raboty s bol'shimi dannymi postroiat v Moskve [A data center for working with big data will be built in Moscow]. Avaliable at: https://www.iksmedia.ru/news/5890168-Datacentr-dlya-raboty-s-bolshimi.html (date of access 09.07.2023). (In Rus)

19. N.V. Nikolaeva, "Big Data: Applications in science and business," I*nformatsionnye tekhnologii v obrazovanii, nauke i promyshlennosti* [Information technologies in education, science and industry]. 2017. No. 2 (16), pp. 81-88. (In Rus)

20. E.V. Zhernakova, A.A. Nikolaev, "Big Data analysis using machine learning technology," *Informatsionnye tekhnologii* [Information technology]. 2018. No. 4, pp. 21-28. (In Rus)

21. A.G. D'iakonov, "Big Data and Machine learning," *Avtomatika i telemekhanika* [Automation and telemechanics]. 2018. No. 79 (10), pp. 1563-1575. (In Rus)

22. A.A. Kamalbekova, A.O. Strekalovskii, "Big Data analysis using machine learning methods," *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naia tekhnika i informatika* [Bulletin of Tomsk State University. Management, computer engineering and computer science]. 2018. No. 44 (4), pp. 48-60. (In Rus)

23. E.S. Patrin, O.U. Kalinova, "Big Data: Statistical Analysis and Machine Learning in economics," *Ekonomika regiona* [Economy of the region]. 2019. No. 15 (3), pp. 736-748. (In Rus)

24. A.I. Poletaev, V.V. Azarov, "Big data and artificial intelligence in the economy," *Vestnik Moskovskogo gosudarstvennogo universiteta im. M. V. Lomonosova. Seriia: Ekonomika* [Bulletin of the Lomonosov Moscow State University. Series: Economics]. 2019. No. 5. Pp. 39-52. (In Rus)

25. I.A. Timoshin, "Big Data and the Internet of Things: analysis and application," *Vestnik Penzenskogo gosudarstvennogo universiteta* [Bulletin of the Penza State University]. 2019. No. 4, pp. 146-154. (In Rus)

26. Big Data has been reassessed. Kommersant. Avaliable at: https://www.kommersant.ru/doc/5939856 (date of access 09.07.2023). (In Rus)

27. Internet traffic (global market). TAdviser. Avaliable at: https://www.tadviser.ru/index.php/ (date of access 09.07.2023). (In Rus)

28. O.A. Kazakovtseva, T.V. Cherkasova, "Big data: problems of information processing and storage," *Izvestiia Ural'skogo gosudarstvennogo universiteta* [Proceedings of the Ural State University]. 2019. No. 187 (3), pp. 68-81. (In Rus)

29. V.I. Luk'ianov, V.I. Riabtsev, "Big data: problems of information analysis and storage," *Vestnik Udmurtskogo universiteta. Seriia: Matematika. Mekhanika. Komp'iuternye nauki* [Bulletin of the Udmurt University. Series: Mathematics. Mechanics. Computer Science]. 2019. No. 29 (2), pp. 220-232. (In Rus)

30. The technology of superstable optical memory was developed in the interests of the Russian Ministry of Defense. Avaliable at: https://function.mil.ru/news_page/country/more.htm?id=12122289@egNews (date of access 10.07.2023)/ (In Rus)

31. K.U. Gusev, T.M. Korablina, E.N. Stepanova, E.A. Zhizheleva, "Application of Hadoop technology for big data processing," *Internet-matematika* [Internet-mathematics]. 2015. T. 11. No. 3, pp. 131-141. (In Rus)

32. T.S. Astapenko, D.D. Sokolin, "Security problems of the Hadoop Big Data processing System," *Reshetnevskie chteniia* [Reshetnev readings]. 2018, pp. 314-315 [Online resource]. Avaliable at: https://cyberleninka.ru/article/n/problemy-bezopasnosti-sistemy-obrabotki-bolshih-dannyh-hadoop (date of access 10.07.2023). (In Rus)

33. Davydova O.V., Savina I.A., Belousova O.I., Petrov A.V., Garifullin R.R. Ispol'zovanie tekhnologii MapReduce dlia obrabotki bol'shikh dannykh [Using MapReduce technology for big data processing]. Avtomatika i telemekhanika [Automation and telemechanics]. 2015. No. 11. Pp. 128-135. (In Rus)

34. A.A. Nekratiuk, O.A. Safar'ian, "Using the MapReduce method in big data," *Molodoi issledovatel' Dona* [Young researcher of the Don]. 2020. no. 3 (24). Avaliable at: https://cyberleninka.ru/article/n/ispolzovanie-metoda-mapreduce-v-big-data (date of access 10.07.2023). (In Rus)

35. O.I. Kukartseva, E.A. Korneeva, V.V. Khramkov, "Actual problems of aviation and cosmonautics," 2022. Vol. 2, pp. 208-209.

36. V.P. Koltakov, "Big data and analytics in information and analytical systems," *Vestnik Moskovskogo universiteta Seriia 15: Vychislitel'naia matematika i kibernetika* [Bulletin of the Moscow University. Series 15: Computational Mathematics and Cybernetics]. 2019. No. 2, pp. 9-17. (In Rus)

37. N.N. Moshak, "Security of information systems," *Saint Petersburg. Saint Petersburg gosudarstvennyi universitet aerokosmicheskogo priborostroeniia* [Saint Petersburg State University of Aerospace Instrumentation]. 2020. 73 p. (In Rus)

38. N.N. Moshak, "Protection of information systems," *Saint Petersburg: Saint Petersburg gosudarstvennyi universitet aerokosmicheskogo priborostroeniia* [Saint Petersburg State University of Aerospace Instrumentation]. 2020. 154 p. (In Rus)

39. N.N. Moshak, L.K. Ptitsyna, "Secure information systems," *Saint Petersburg. Saint Petersburg gosudarstvennyi universitet Telekommunikatsii im. professora M. A. Bonch-Bruevicha* [Saint-Petersburg State University of Telecommunications named after Professor M. A. Bonch-Bruevich]. 2020. 216 p. (In Rus)

40. N.N. Moshak, "Fundamentals of Information security management," *Saint Petersburg. Saint Petersburg gosudarstvennyi universitet aerokosmicheskogo priborostroeniia* [Saint Petersburg State University of Aerospace Instrumentation]. 2022. 141 p. (In Rus)

41. Zashchita Big Data: problemy i resheniia [Big Data Protection: problems and solutions]. [Online resource]. Avaliable at: https://www.itworld.ru/cionews/security/158148.html (date of access 10.07.2023). (In Rus)

42. Giuliara Muradova, "Issues of information security of "big data"," II Republican scientific-practical conference on multidisciplinary problems of Information Security. 2015, pp. 228-231. Avaliable at: https://ict.az/uploads/ konfrans/2_konfrans/63.pdf (date of access 11.07.2023). (In Rus)

INFORMATION ABOUT AUTHORS:

Nikolai N. Moshak, Dr.Sci.Tech., professor of St. Petersburg State University of Aerospace Instrumentation st, St. Petersburg, Russia;
Dr.Sci.Tech., professor of Saint Petersburg State University of Telecommunications of the prof. M.A. Bonch-Bruyevich, St. Petersburg, Russia, nnmoshak49@mail.ru

Sabina R. Rudinskaya, Senior lecturer, Educational Institution "Belarusian State Academy of Communications," Republic of Belarus, Minsk, Republic of Belarus, sabina.rudin@mail.ru

Alexey A. Gruzdev, student of Saint Petersburg State University of Telecommunications of the prof. M.A. Bonch-Bruyevich, St. Petersburg, Russia, gruzdev.a.a26@mail.ru

For citation: Moshak N.N., Rudinskaya S.R., Gruzdev A.A. Third platform for informatization and big data. H&ES Reserch. 2023. Vol. 15. No 4. P. 47-59. doi: 10.36724/2409-5419-2023-15-4-47-59 (In Rus)

doi: 10.36724/2409-5419-2023-15-4-60-67

ПРОГРАММНОЕ СРЕДСТВО, ОПРЕДЕЛЯЮЩЕЕ ФЕЙКОВЫЙ ВИДЕОКОНТЕНТ С ПОМОЩЬЮ ТЕХНОЛОГИИ DEEPFAKE АЛГОРИТМА GAN

ДЖУРОВ
Александр Андреевич[1]

ЧЕРКЕСОВА
Лариса Владимировна[2]

РЕВЯКИНА
Елена Александровна[3]

Сведения об авторах:

[1]аспирант профиля "Информационные системы и процессы", кафедра "Кибербезопасность информационных систем", Донской Государственный Технический Университет (ДГТУ), г. Ростов-на-Дону, Россия
sashaz1696@yandex.ru

[2]доктор физико-математических наук, профессор кафедры "Кибербезопасность информационных систем", факультет "Информатика и Вычислительная техника", акад. Российской Академии Естествознания, чл.-корр. Международной Академии Наук Прикладной Радиоэлектроники, чл.-корр. Российской Академии Изучения Проблем Национальной Безопасности, Донской Государственный Технический Университет (ДГТУ), г. Ростов-на-Дону, Россия, chia2002@inbox.ru

[3] доцент, кандидат технических наук кафедры "Кибербезопасность информационных систем", факультет "Информатика и Вычислительная техника", Донской Государственный Технический Университет (ДГТУ), г. Ростов-на-Дону, Россия, revyelena@yandex.ru

АННОТАЦИЯ

Введение: в современном мире одним из основных и актуальных проблем является ложный контент: новости, видео, фото и тд. На раннем этапе развития технологии Deepfake, она использовалась пользователями-любителями для генерации мультимедийного контента путем сопоставления человеческих выражений лиц и фраз, "хозяевами" которых, как правило являлись узнаваемые личности, для создания фейковых СМИ, выглядящих подлинными. Но ситуация меняется, и технология Deepfake начинает использоваться не для компрометации, а для агитации и привлечения политических сторонников. **Цель исследования:** Программная реализация алгоритма распознавания видеоконтента, синтезированного с помощью технологии Deepfake алгоритма GAN, с приемлемой точностью. В работе была предложена программная реализация, которая анализирует видео и выводит решение о подлинности данного. Представлены основные архитектуры алгоритма GAN, а возможности и угрозы применения технологии deepfake. Проведен анализ особенностей моделей Xception и ResNeXt, обученных с помощью нейронных сетей. Методы: Для работы системы необходимо осуществить выбор подходящих нейронных сетей на основе результатов производительности, которыми могут быть ResNeXt, XceptionNet или любая другая нейронная сеть. В рамках данной работы будут рассмотрены и использованы в программной реализации именно ResNeXt и XceptionNet, а также BlazeFace является предварительно обученной моделью распознавания человеческих лиц, используется для распознавания лиц на извлеченных изображениях. **Результаты:** На вход функции подается путь к видео (в файловой системе). Образец проходит покадровую проверку на наличие лица в каждом отдельном фрейме, если распознание прошло успешно, данные добавляются в список. По желанию можно оставить фиксированное количество семплов с наилучшим качеством среди представленных.

КЛЮЧЕВЫЕ СЛОВА: *Deepfake, GAN, нейронная сеть, информационная безопасность, дискриминатор.*

Для цитирования: Джуров А.А., Черкесова Л.В., Ревякина Е.А. Программное средство, определяющее фейковый видеоконтент с помощью технологии deepfake алгоритма GAN // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 60-67. doi: 10.36724/2409-5419-2023-15-4-60-67

Введение

Технология Deepfake [1] – это методика компьютерного синтеза изображения, основанная на искусственном интеллекте, которая используется для соединения и наложения существующих изображений и видео на исходные изображения или видеоролики. Искусственный интеллект использует синтез изображения человека – объединяет несколько картинок, на которых человек запечатлен с разных ракурсов и с разным выражением лица, и делает из них видео [2].

Deepfake представляет собой данные, полученные с помощью синтеза, в содержании которых личность и его лицо из реального видеоряда заменяется на другую личность. Как правило, результаты имеют формат аудиозаписи, фото или видео (на данный момент).

С течением некоторого времени ходовые генеративные модели начали показывать многообещающие результаты в создании реалистичных изображений. Результаты разработки Гудфеллоу показали большие успехи также и в области компьютерного зрения (КЗ). В последнее время он начал показывать перспективные результаты как в генерации высококачественного аудио, так и видеоконтента [3].

Еще одна угроза, которая может усилиться при использовании дипфейков – это дезинформация в политике. Дипфейки могут быть быстро созданы и легко распространены среди широкой аудитории.

Обладая этим специфическим преимуществом, дипфейк может быть сознательно или неосознанно использован для дезинформации общественности в политических целях. Например, использование дипфейкового видео итальянским сатирическим телешоу против официального премьер-министра Италии Маттео Ренци. На видео, опубликованном в социальных сетях, изображено, как он оскорбляет коллег-политиков. Когда видео распространилось в сети, многие люди начали верить, что видео было подлинным, что вызвало возмущение общественности [4].

Алгоритм GAM

В алгоритме GAN, популярность которого возрастает, используются искусственные нейронные сети (ИНС). В терминологии сферы ИИ они именуются как синтезатор (генератор) и дискриминатор (детектор) [5].

Генерирующий алгоритм, на вход которому поступают случайные данные, синтезирует уникальный контент. Другая ИНС, являющаяся дискриминатором, проверяет контент, чтобы убедиться, что они соответствует исходникам. Данная конкуренция двух ИНС, по сути, и составляет основной принцип работы GAN, ИНС-синтезатор на выходе преподносит реалистичные данные, в том числе с лицами известных людей. Рассматривая это через призму математических вычислений, нейросети, синтезирующие картинки (статика) и видео (динамика) не имеют различий, при том, что они могут использоваться для разных вещей.

В процессе создания фейкового видеоряда генерируется множество последовательных изображений, это обусловлено тем, что есть необходимость придавать движениям людей плавность, чтобы избегать резкого движения частей тела от кадра к кадру. Эта самая плавность достигается за счёт различный модификация алгоритма GAN, которые помимо всего, учитывают то, что было на предыдущих итерациях [6].

Чтобы улучшить 3D-изображение объекта на видео, в нейронную сеть необходимо загрузить фотографии объекта, сделанные с разных ракурсов. Если вы будете одинаково фотографировать людей с бородой и без бороды, вы не получите точных результатов. Поэтому не стоит бояться, что злоумышленники будут брать фотографии из социальных сетей и создавать дипфейки на основе ваших изображений [7].

Для того чтобы создать качественное искусственное изображение на основе фотографий, придется сделать несколько снимков, снятых с разных ракурсов, вручную создать 3D-модель, синтезировать множество отдельных изображений этой 3D-модели и загрузить их в нейросеть [8].

Этап 1 – происходит определение черт лица в картинках (кадрах), имеющихся из искомого видео. Затем, для прочего упрощения сложных вычислений, происходит отбрасывание некорректных (неудачных) кадров.

Этап 2 – происходит процесс определения контуров человеческого лица в картинках, полученных из второго видеоряда. Главным отличием является то, что в данной ситуации необходимо извлечь все лица в каждом отдельном образце, даже если лицо будет нечетким или мутным.

Этап 3 – обучение ИНС на полученных датасетах с изображениями и видеорядами. Для тренировки необходимо выбрать одну из моделей обучения, подобрать архитектуру. Тренировка в сою очередь является базовой циклической процедурой, выполняемой ИНС относительно алгоритма GAN. От качества данных для обучения зависит и качество работы ИНС.

По результатам обучения на 4 этапе производится покадровое наложение сгенерированных лиц на изображения, полученные из исходного материала. Возможно использование нескольких режимов наложения.

Этап 5 – конечная стадия, процесс наложения кадров в видео с ровно той же частотой фреймов, звуковым сопровождением, что в исходнике [9].

Каждая стадия работы процесса требует различных временных ресурсов как от человека и от ЭВМ. Время работы программного средства, покадрово извлекающего изображения из видео, может составить несколько минут, однако для проверки результатов человеку могут потребоваться часы.

Вот некоторые из самых популярных направлений GAN, которые активнее всего рассматриваются научным сообществом: конвертация исходной картинки между состояниями (CycleGAN), создание изображения на основе текстового описания (преобразование текста в изображение), напечатанного или даже написанного от руки человека, а после распознанное ИИ, создавание изображения с очень высоким разрешением (развитие классического алгоритма до идеала).

Система состоит из двух нейросетей – генератора и дискриминатора (детектора), которые обучаются по методу backpropagation (метод обратного распространения ошибки). Суть метода строится на том, что распространение сигналов об неточности в значениях искомых входных и выходных точек (input-output), в направлении, негативном прямому распространения в стандарте [10].

На рисунке 1 представлена схема работы алгоритма GAN.

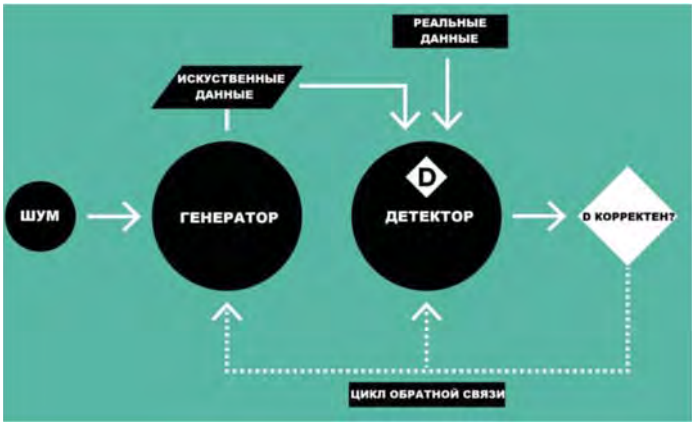


Рис. 1. Схема работы алгоритма GAN

Генератор создает из множества случайных чисел (случайного шума из заранее выбранного распределения) картинку, причем изображение должно быть максимально реалистичным. Синтез происходит на основе имеющегося набора данных. Далее данные передаются детектору.

В дискриминатор (детектор) попадает образец из генератора, а также искомое изображение. Представляет собой двоичный классификатор, который пытается с наибольшей точностью определить, является ли входная выборка реальной (вывод скалярного значения 1) или ложной (вывод скалярного значения 0). Причем в генератор поступает информация о том, по какой причине дискриминатор определил текущую выборку как синтезированный контент.

Дискриминатор хочет выполнять свою работу максимально качественно. Когда поддельный образец (созданный генератором) передается дискриминатору и производятся соответствующие вычисления, результаты всегда округляются не в пользу генератора, но последний хочет сгенерировать образцы таким образом, чтобы дискриминатор допустил ошибку, назвав его подлинным.

В конце каждой итерации детектор получает информацию от специального блока, правильно ли он выполнил свою работу или нет. Данный блок называется блоком потерь или функцией Loss.

Основные методы и средства работы ПО

При разработке программного средства используются следующие модули Python:

– openCV – библиотека компьютерного зрения, которая предназначена для анализа, классификации и обработки изображений [11];

– NumPy – это открытая бесплатная Python-библиотека для работы с многомерными массивами [12];

– Pandas – высокоуровневая Python-библиотека для анализа данных [13];

– PyTorch – современная библиотека глубокого обучения [14];

– Deepfakeutils – библиотека, включающая в себя модели обучения, а также инструменты, необходимые для создания и детекции Deepfake-контента.

В разработанном программного обеспечения использованы в реализации именно ResNeXt [15] и XceptionNet [16]. Архитектура Xception (рис. 2) основана на теории, что обработка двух типов информации непосредственно в последовательности не приводит к снижению качества сети, и разлагает традиционную свертку на кросс-канальную (которая имеет дело только с межканальными корреляциями) и пространственную (которая имеет дело только с пространственными корреляциями внутри каждого канала). Получившаяся на рисунке конструкция и составляет полный модуль Inception [17]. Архитектура Inception, предложенная группой разработчиков в 2015 году, не выбирает размер ядра, а использует несколько массивов одновременно, которые восстанавливаются в одно и то же время, и использует слияние для вывода каналов.

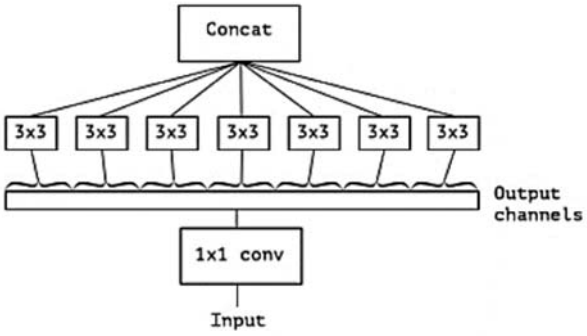


Рис. 2. Схема блока Xception

Вместо обычного алгоритма блока нейронной сети выполняется последовательно два шага. 1 Свернём имеющийся тензор размером 1x1 свёрткой, подобно тому, как это выполнялось в блоке Inception, получив новый тензор. Эта операция называется pointwise convolution.

2 Свернём каждый канал по отдельности сверткой с параметрами 3 × 3 (в этом случае размерность не изменится, так как мы сворачиваем не все каналы вместе, как в обычном сверточном слое). Эта операция называется depthwise spatial convolution.

Модели нейронных сетей обучались с использованием выбранного алгоритма нейронной сети на сервисе Google Cloud Platform, используемом разработчиками в сфере искусственного интеллекта и машинного обучения.

Сверточный слой в данной архитектуре обрабатывает внутриканальную и межканальную информацию последовательно, но в рамках одного процесса. Это позволяет существенно снижать нагрузку, так как количество весов в рамках одного вычисления будет снижено. На рисунке 3 изображена схема работы блока ResNeXt.

Данный метод основывается на том, что мощность (размер серии преобразований) – это конкретное измеримое значение (не константа), которое имеет центральное значение наряду с измерениями ширины и глубины.

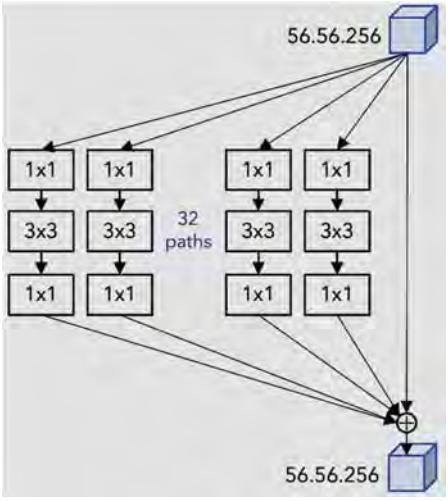


Рис. 3. Схема блока ResNeXt

Эксперименты показывают, что повышение производительности является более эффективным способом повышения точности, чем углубление или расширение, особенно когда глубина и ширина начинают давать существующим моделям меньшие результаты при анализе функции потерь.

Простейшие нейроны в искусственных нейронных сетях выполняют внутреннее произведение (взвешенную сумму), которое представляет собой элементарное преобразование, выполняемое полносвязными и сверточными слоями [18].

Соответственно, чем больше вес внутри отдельно взятого слоя, тем больше вероятность, что характерное его признаки будут доминирующими при обучении и при дальнейшем распределении весов. Данная серия вычислений называется wixi. Эта операция отображена на рисунке 4.

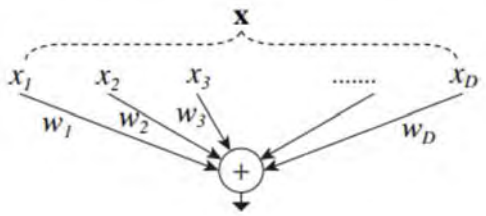


Рис. 4. Серия вычислений wixi

Схема работы разработанного ПО

Основной алгоритм проверки видео на подделку состоит из следующих обязательных шагов.

1. Загрузка пред обученных моделей ResNeXt и Xception, делается это с помощью модуля gdown, одного из базовых библиотек по скачиванию файлов из Интернета.

2. Переход в модуль получения результатов предугадывания относительно модели ResNeXt. Выводом является величина отклонения в предугадывании.

3. Переход в модуль получения результатов предугадывания относительно модели Xception. Выводом является величина отклонения в предугадывании.

4. Вычисление среднего значения точности предугадывания, находится среднее арифметическое коэффициентов отклонений.

5. Если среднего значения точности предугадывания больше порогового значения (задается вручную), то выводом является заключение о том, что видео является фейковым. Иначе – выводом является заключение о его реальности.

Данные шаги является действенным способом по определению фейкового видеоконтента, синтезированного с помощью технологии Deepfake алгоритма GAN, работающим с приемлемой точностью. Описанный алгоритм так же представлен в качестве блок-схемы на рисунке 5.

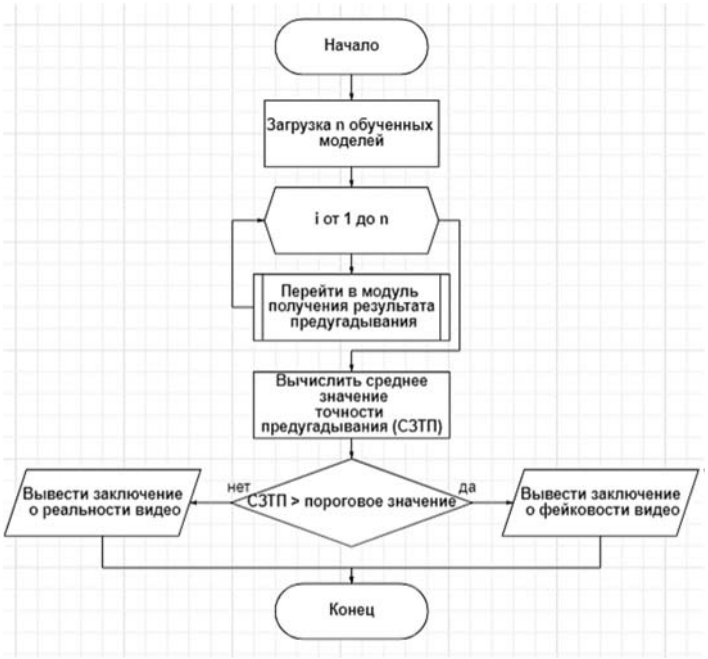


Рис. 5. Блок-схема работы основного алгоритма программы

Также в данной программе модуль получения результатов предугадывания отдельной модели. Алгоритм состоит из следующих шагов.

1. Извлечение кадров с лицами из искомого видеоряда.

2. Для каждого кадра происходит его форматирование для конкретной модели.

3. Нормализация данных каждого кадра. Подача кадров на анализ. Получение результатов анализа.

4. Вывод результата анализа (в консоль или в качестве видеоряда).

Алгоритм представлен в качестве блок-схемы на рисунке 6.

Помимо этого, программная реализация включает в себя модуль анализа данных. Его работа состоит из нескольких этапов.

1. Преобразование данных каждого кадра в тензор.

2. Если модель Xception, то анализ происходит относительно этой модели, иначе – относительно ResNeXt.

3. Отключение градиентного спуска.

4. Выгрузка модели. Запуск процесса предугадывания.

5. Нормализация данных с помощью сигмоиды.

6. Вывод среднего значения всех элементов массива.

Рис. 6. Блок-схема алгоритма получения результата предугадывания

На рисунке 7 представлена блок-схема работы данного модуля.

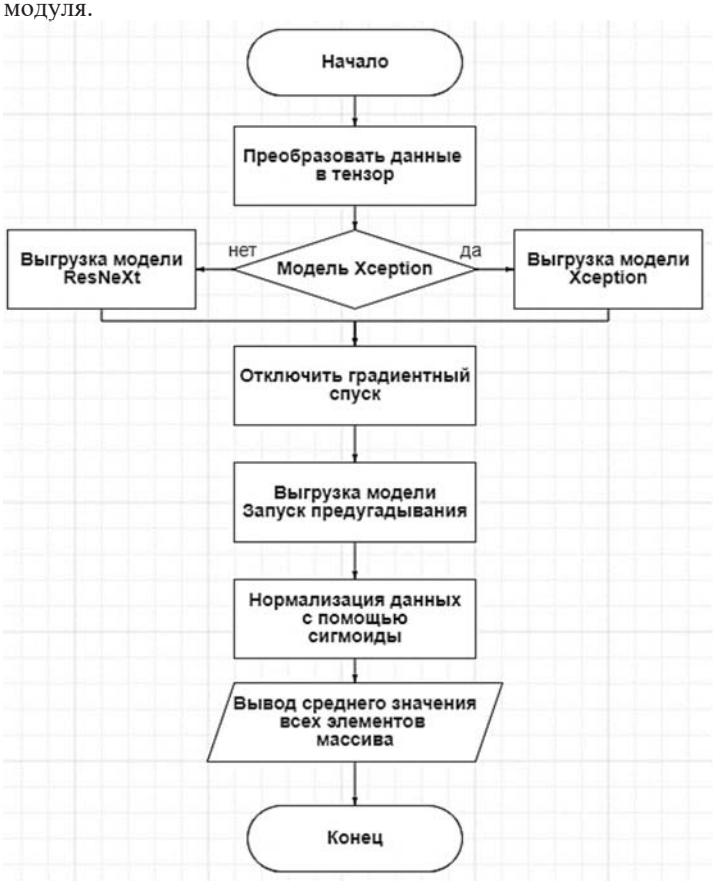


Рис. 7. Блок-схема алгоритма анализа данных

Принцип работы программного средства

Программа принимает на вход строку, которая является путем к видео в файловой системе, проверяет корректность введённых данных, а именно отсутствие повторяющихся функций, отсутствие взаимоисключающих аргументов и отсутствие неверных опций.

Затем программа начинает свою работу на любом незагруженном логическом ядре, в противном случае будет выведено сообщение о возможном снижении производительности процесса тестирования. Далее проверяется возможность создания выходных файлов и чтение тестовых данных. Сбои во время этих действий маловероятны, однако может произойти ситуация, когда необходимых файлов нет. Программа не имеет графического интерфейса, поэтому запуск осуществляется через консоль. На рисунке 8 представлен пример запуска.

Если всё отработало удачно, тогда запускается тестируемая программа для всех тестовых данных, чтобы убедиться, что она работает должным образом. Это делается только для начальных входных данных и только один раз.

```
PS F:\Programm> & C:/Users/roone/AppData/Local/Programs/Python/Python39/python.exe f:/Programm/detector.py

Choose what you want:
1 - Get text result about one video
2 - Get text result about few video
3 - Get video result about one video
Mode: 1

Welcome to "Deeptector"! Start working!

Choose video to analyze:
        1 - examples\abbalbisk.mp4
        2 - examples\capitain_lore.mp4
        3 - examples\donald_thrump.mp4
        4 - examples\krylova.mp4
        5 - examples\me_at_zoo.mp4
        6 - examples\morgan_freeman.mp4
        7 - examples\queen_elizabeth.mp4
        8 - examples\robert_downey.mp4
        9 - examples\vladimir_putin.mp4
        10 - examples\vladimir_zelenskiy.mp4
Enter the number: 1

Load Xception pre-trained model...
        Done!

Load ResNext pre-trained model...
        Done!

Face samples:
        64

Model prediction:
        0.11014726758003235

Face samples:
        64

Model prediction:
        0.07376221567392349

Comparing prediction and threshold:
        0.08350665807823768 < 0.3

examples\abbalbisk.mp4 - REAL!

Good Bye!
```

Рис. 8. Работа программы в первом режиме

Разберем аргументы. В данном случае основными входными параметрами являются пути к видео, которые могут быть представлены как строка, объект Path библиотеки pathlib или объект Windows Path библиотеки os. Для работы программы необходимо ввести номер режима, первый режим представляет собой вывод заключения о поддельности видео

в виде текстовой строки, второй режим выводит строку с заключением о фейковости произвольного количества образцов, при работе третьего режима пользователь получает видео-заключение с вердиктом программы.

На рисунке 9 представлен пример работы второго режима программы.

```
Mode: 2
----------
Welcome to "Deeptector"! Start working!
----------
Choose videos you want to analyze:
        1 - examples\abbalbisk.mp4
        2 - examples\capitain_lore.mp4
        3 - examples\donald_thrump.mp4
        4 - examples\krylova.mp4
        5 - examples\me_at_zoo.mp4
        6 - examples\morgan_freeman.mp4
        7 - examples\queen_elizabeth.mp4
        8 - examples\robert_downey.mp4
        9 - examples\vladimir_putin.mp4
        10 - examples\vladimir_zelenskiy.mp4
Choose number (press q to exit): 1
Choose number (press q to exit): 3
Choose number (press q to exit): 7
Choose number (press q to exit): q
----------
Face samples:
        64
----------
Model prediction:
        0.3829379081726074
----------
Face samples:
        64
----------
Model prediction:
        0.4184581935405731
----------
Face samples:
        64
----------
Model prediction:
        0.7207198143005371
----------
[0.3829379081726074, 0.4184581935405731, 0.7207198143005371]
----------
Comparing prediction (ResNeXt) and threshold:
        0.3829379081726074 > 0.3
----------
Comparing prediction (ResNeXt) and threshold:
        0.4184581935405731 > 0.3
----------
Comparing prediction (ResNeXt) and threshold:
        0.7207198143005371 > 0.3
----------
abbalbisk.mp4 - FAKE!
donald_thrump.mp4 - FAKE!
queen_elizabeth.mp4 - FAKE!
----------
Good Bye!
```

Рис. 9. Работа программы во втором режиме

Видео, которые были проанализированы детектором являются дипфейками и оригинальными видео разного уровня качества, как самого видео, так и степени точности подделывания. Из результатов запуска видно, что детектор ошибся в одном случае из трех (рис. 10).

В среднем проверка одного видео занимает от семи до тридцати секунд в зависимости от длительности контента, вся тестовая сессия длиться в среднем девяносто секунд.

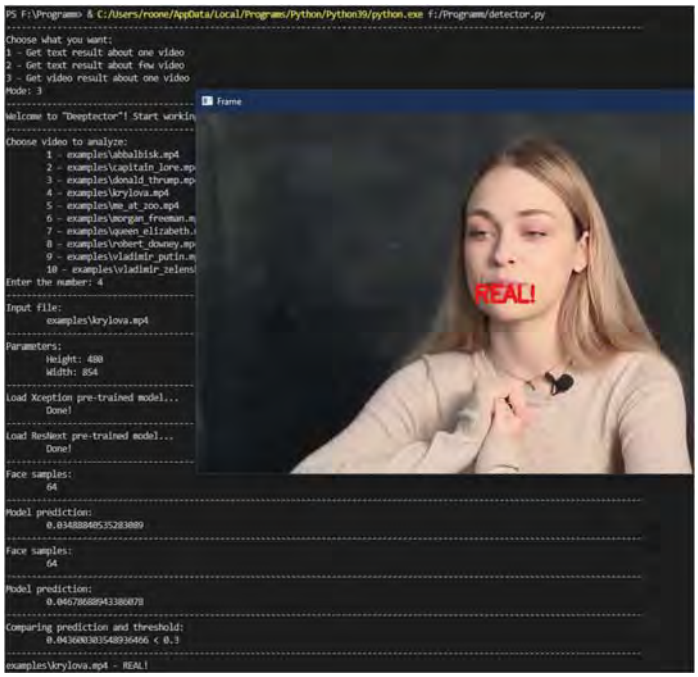


Рис. 10. Демонстрация работы программы в третьем режиме

Третий режим работает идентично первому, основное отличие заключается в том, что в конце выводится результирующее видео. В данном режиме вводиться дополнительный аргумент, который представляет собой имя файла-результата. Пример работы программы приведен на рисунке.

Программное средство принимает на вход тестовые данные, обрабатывает их, выводит итоги тестирования и генерировать файлы-результаты с вердиктом о синтезированности искомого видеоряда.

Заключение

Технологии глубокого подражания быстро развиваются. Точность получаемых данных растет. Алгоритмы обнаружения мошенничества совершенствуются. Предполагается, что в будущем многие компании будут предлагать услуги по борьбе с данной угрозой [19].

В настоящее время двумя факторами, препятствующими широкому распространению дипфейков, являются низкий уровень совершенства алгоритмов и высокая стоимость конечного продукта. Однако с ростом индустрии развлечений эти два показателя будут улучшаться по мере выхода deepfake на массовый рынок [20].

Было реализовано программное средство для проверки видео на предмет его синтезированности с помощью технологии Deepfake алгоритма генеративно-состязательных сетей (GAN) на языке Python. Хранение значений кадров с лицами и соответствующих им имен файлов происходит в оперативной памяти. Проведено тестирование программного средства для различных файлов, включая обработку исключительных ситуаций.

Литература

1. *Барабанщиков В.А.* Deepfake в исследованиях восприятия лица. М.: Издательство ИНГН, 2018. 176 с.
2. *Баранова Е.К.* Информационная безопасность и защита информации с нуля до полного понимания. М.: Риор, 2018. 400 с.
3. *Lyu S.* Deepfake Detection: Current Challenges and Next Steps // IEEE International Conference on Multimedia & Expo Workshops (ICMEW). 2020. P. 1-6. DOI: 10.1109/ICMEW46912.2020.9105991.
4. *Xinyi Z., Reza Z.* A Survey of Fake News: Fundamental Theories, Detection Methods and Opportunities // ACM Computing Surveys. 2020. Vol. 53, Iss. 5, pp. 1-40. DOI: 10.1145/3395046.
5. *Dash A., Ye J., Wang G.* A review of Generative Adversarial Networks (GANs) and its applications in a wide variety of disciplines // arXiv preprint. 2021.
6. *Ярочкин В.И.* Информационная безопасность. М.: Академический проект, 2018. 544 с.
7. *Крон Д.* Глубокое обучение в картинках. Визуальный гид по ИИ. Санкт-Петербург: Питер, 2016. 416 с.
8. *Стюарт Р.* Искусственный интеллект. Современный подход к решению актуальной проблемы. М.: МГИУ, 2017. 272 с.
9. *Курцвейл Р.* Как создать разум: секрет человеческого мышления раскрыт. Санкт-Петербург: BHV, 2019. 368 с.
10. *Songyuan L., Fan M., Chen R.* Overview of generative adversarial networks. J Phys Conf Ser. 2021.
11. *Howse J., Minichino J.* Learning OpenCV 4 Computer Vision with Python 3: Get to grips with tools, techniques, and algorithms for computer vision and machine learning. 3rd Editio, 2020. 372 p.
12. *Johansson R.* Numerical Python: Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib. 2019. DOI: 10.1007/978-1-4842-4246-9.
13. *Лемешевский С.В.* Введение в библиотеку pandas. Институт математики НАН Беларуси. 2020.
14. *Макмахан Б., Рао Д.* Знакомство с PyTorch. Глубокое обучение при обработке естественного языка. 2020. ISBN:978-5-4461-1241-8.
15. *Zhou T., Zhao Y., Wu J.* Resnext and res2net structures for speaker verification. Microsoft Corporation, USA. 2020.
16. *Jain A.K.* Artificial neural networks: a tutorial // Computer. 1996. Vol. 29. № 3, pp. 31-44. DOI: 10.1109/2.485891.
17. *Do N.Q.* Phishing webpage classification via deep learning-based algorithms: an empirical study// Applied Sciences. 2021. Vol. 11. № 19. 32 p. DOI: 10.3390/app11199210.
18. *Letou K.* Host-based Intrusion Detection and Prevention System (HIDPS) // International Journal of Computer Applications. 2013. Vol. 69. № , pp. 28-33. DOI: 10.5120/12136-8419.
19. *Mitchell T.M.* Machine learning. New York: McGraw-hill. 1997. Vol. 1. № 9. 414 p. ISBN 0071154671, 9780071154673.
20. *Goodfellow I.* Deep learning. MIT press. 2016. 800 p. ISBN 9780262337373, 0262337371.

SOFTWARE TOOL FOR DETECTING FAKE VIDEO CONTENT USING THE DEEPFAKE TECHNOLOGY OF THE GAN ALGORITHM

ALEXANDER A. DZHUROV
Rostov-on-Don, Russia

LARISA V. CHERKESOVA
Rostov-on-Don, Russia

ELENA A. REVYAKINA
Rostov-on-Don, Russia

KEYWORDS: *Deepfake, GAN, neural network, Information Security, discriminator.*

ABSTRACT

Intoduction: in the modern world, one of the main and urgent problems is false content: news, videos, photos, etc. Early on in the development of Deepfake technology, it was used by amateur users to generate multimedia content by matching human facial expressions and phrases, usually owned by recognizable individuals, to create fake media that looked genuine. But the situation is changing, and Deepfake technology is being used not for compromising, but for campaigning and attracting political supporters. **The purpose of the study:** Software implementation of the video content recognition algorithm, synthesized using the Deepfake technology of the GAN algorithm, with acceptable accuracy. In the work, a software implementation was proposed that analyzes the video and makes a decision about the authenticity of this one. The main architectures of the GAN algorithm are presented, as well as the opportunities and threats of using deepfake technology. An analysis of the features of the Xception and ResNeXt models trained using neural networks was carried out. **Methods:** For the system to work, it is necessary to select suitable neural networks based on performance results, which can be ResNeXt, XceptionNet or any other neural network. As part of this work, ResNeXt and XceptionNet will be considered and used in the software implementation, as well as BlazeFace is a pre-trained human face recognition model used to recognize faces in extracted images. **Results:** The function input is the path to the video (in the file system). The sample is frame-by-frame checked for the presence of a face in each individual frame, if the recognition was successful, the data is added to the list. Optionally, you can leave a fixed number of samples with the best quality among those presented.

REFERENCES

1. V.A. Drummers, “Deepfake in face perception research,” Moscow: INGN Publishing House, 2018. 176 p. (In Rus)

2. E.K. Baranova, Information security and information protection from scratch to full understanding,” Moscow: Rior, 2018. 400 p. (In Rus)

3. S. Lyu, “Deepfake Detection: Current Challenges and Next Steps,” *IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. 2020, pp. 1-6. DOI: 10.1109/ICMEW46912.2020.9105991.

4. Z. Xinyi, Z. Reza, “A Survey of Fake News: Fundamental Theories, Detection Methods and Opportunities,” *ACM Computing Surveys*. 2020. Vol. 53, Iss. 5, pp. 1-40. DOI: 10.1145/3395046.

5. A. Dash, J. Ye, G. Wang, “A review of Generative Adversarial Networks (GANs) and its applications in a wide variety of disciplines,” arXiv preprint. 2021.

6. V.I. Yarochkin, “Information Security,” Moscow: Academic project, 2018. 544 p. (In Rus)

7. D. Kron, “Deep learning in pictures,” Visual guide to AI. St. Petersburg: Piter, 2016. 416 p. (In Rus)

8. R. Stuart, “Artificial intelligence. A modern approach to solving an actual problem,” Moscow: MGIU, 2017. 272 p. (In Rus)

9. R. Kurzweil, “How to create a mind: the secret of human thinking is revealed,” St. Petersburg: BHV, 2019. 368 p. (In Rus)

10. L. Songyuan, M. Fan, R. Chen, “Overview of generative adversarial networks,” *J Phys Conf Ser.* 2021.

11. J. Howse, J. Minichino, “Learning OpenCV 4 Computer Vision with Python 3: Get to grips with tools, techniques, and algorithms for computer vision and machine learning,” 3rd Edition, 2020. 372 p.

12. R. Johansson, “Numerical Python: Scientific Computing and Data Science Applications with Numpy,” *SciPy and Matplotlib*. 2019. DOI: 10.1007/978-1-4842-4246-9.

13. S.V. Lemeshevsky, “Introduction to the pandas library,” Institute of Mathematics of the National Academy of Sciences of Belarus. 2020. (In Rus)

14. B McMahan, D. Rao, “Introduction to PyTorch,” *Deep learning in natural language processing*. 2020. ISBN:978-5-4461-1241-8. (In Rus)

15. T. Zhou, Y. Zhao, J. Wu, “Resnext and res2net structures for speaker verification,” Microsoft Corp., USA. 2020.

16. A.K. Jain, “Artificial neural networks: a tutorial,” *Computer.* 1996 Vol. 29. No. 3, pp. 31-44. DOI: 10.1109/2.485891.

17. N.Q. Do, “Phishing webpage classification via deep learning-based algorithms: an empirical study,” *Applied Sciences*. 2021. Vol. 11. No. 19. 32 p. DOI: 10.3390/app11199210.

18. K. Letou, “Host-based Intrusion Detection and Prevention System (HIDPS),” *International Journal of Computer Applications.* 2013. Vol. 69. No. 26, pp. 28-33. DOI: 10.5120/12136-8419.

19. T.M. Mitchell, “Machine learning,” New York: McGraw-hill. 1997 Vol. 1. No. 9. 414 p. ISBN 0071154671, 9780071154673.

20. I. Goodfellow, “Deep learning,” *MIT press*. 2016. 800 p. ISBN 9780262337373, 0262337371.

INFORMATION ABOUT AUTHORS:

Alexander A. Dzhurov, postgraduate student of the profile "Information systems and processes", Department of Cybersecurity of information systems, Don State Technical University (DSTU), Rostov-on-Don, Russia, sashaz1696@yandex.ru

Larisa V. Cherkesova, Doctor of Physical and Mathematical Sciences, Full Professor of the Department of Cyber Security of Information Systems, Faculty of Informatics and Computer Engineering, acad. Russian Academy of Natural Sciences, corresponding member. International Academy of Sciences of Applied Radioelectronics, corresponding member. Russian Academy for the Study of National Security Problems, Don State Technical University (DSTU), Rostov-on-Don, Russia, chia2002@inbox.ru

Elena A. Revyakina, Assistant professor, PhD, Department of Cybersecurity of Information Systems, Faculty of Informatics and Computer Science, Don State Technical University (DSTU), Rostov-on-Don, Russia, revyelena@yandex.ru

For citation: Dzhurov A.A., Cherkesova L.V., Revyakina E.A. Software tool for detecting fake video content using the Deepfake technology of the GAN algorithm. *H&ES Reserch*. 2023. Vol. 15. No 4. P. 60-67. doi: 10.36724/2409-5419-2023-15-4-60-67 (In Rus)

doi: 10.36724/2409-5419-2023-15-4-68-74

ДИСТАНЦИОННЫЙ МОНИТОРИНГ СОБЫТИЙ, ВЫЗЫВАЮЩИХ СНИЖЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ ИНФОРМАЦИОННЫХ И ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ

МИХАЙЛОВ
Владимир Юрьевич[1]

МАЗЕПА
Роман Богданович[2]

ВАКУЛЬЧИК
Ольга Витальевна[3]

АННОТАЦИЯ

Введение: Объектом исследования являются события, случайно или намеренно эксплуатирующие встроенный механизм блокировок SQL Server, которые существенно снижают производительность информационных и информационных поисковых систем. Предмет исследования: дистанционный мониторинг блокирующих процессов. **Цель исследования:** разработка защищенного метода дистанционного мониторинга, способного оперативно и надежно обнаружить события, вызывающие эксплуатацию встроенного механизма блокировок SQL Server. **Методы:** Применение мониторинга необходимо для поддержания требуемого уровня производительности системы, а также для оперативного обнаружения воздействий на информационную систему, разрушающих целостность баз данных. Данная статья нацелена на решение основных проблем дистанционного мониторинга: обеспечение полноты оценки состояния и надёжную, оперативную передачу полученных сведений по открытому каналу. Ключевой особенностью предложенного метода дистанционного мониторинга является выявление объектов SQL Server и конструкций языка SQL, применяемых для реализации дистанционного мониторинга с учетом полноты получаемых сведений. Высокая степень безопасности информационной среды дистанционного мониторинга обеспечивается системными средствами реализации авторской клиент-серверной архитектуры. **Результаты:** показано, что оперативность и безопасность передачи ценных управляющих данных обеспечивается применением хранимых процедур. Известные хранимые процедуры избыточны, поэтому авторам потребовалось создать собственный инструмент мониторинга, что и является главным результатом данной статьи. Результаты исследования могут быть применены при проектировании надёжных и безопасных информационных и информационно поисковых систем различного назначения.

Сведения об авторах:

[1] д.т.н., Московский авиационный институт (национальный исследовательский университет), Москва, Россия, mihvj@yandex.ru

[2] к.т.н., Московский авиационный институт (национальный исследовательский университет), Москва, Россия, mrb402@mail.ru

[3] Московский авиационный институт (национальный исследовательский университет), Москва, Россия, olga.vakulchik@mail.ru

КЛЮЧЕВЫЕ СЛОВА: *дистанционный мониторинг состояния, удаленное администрирование, целостность данных, механизм блокировок SQL Server, SQL-запрос.*

Для цитирования: *Михайлов В.Ю., Мазепа Р.Б., Вакульчик О.В.* Дистанционный мониторинг событий, вызывающих снижение производительности информационных и информационно-поисковых систем // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 4. С. 68-74. doi: 10.36724/2409-5419-2023-15-4-68-74

Введение

Автоматизация трудоёмкого процесса хранения, обработки и оперативного предоставления большого объема данных во многих направлениях информационной деятельности базируется на применении электронных баз данных [1]. В совокупности они представляют собой электронное хранилище структурированных данных – базу данных и систему управления базами данных (СУБД), обеспечивающую надежный доступ и оперативное управление данными посредством специальных запросов на языке SQL [2]. В дальнейшем, под базой данных (БД) будем понимать совокупность непосредственно хранилища данных и СУБД.

Корректное и безопасное функционирование информационных систем, использующих базы данных, должно обеспечиваться администраторами и специалистами по информационной безопасности (ИБ). Это специалисты высокой квалификации, потребность в них высока, поэтому не всегда может быть удовлетворена. Тем не менее, некоторые аспекты администрирования, такие как мониторинг состояния сервера [3], могут быть успешно автоматизированы. Одним из направлений такой автоматизации является дистанционный мониторинг состояния удаленной информационной системы и БД. Однако, на этом пути возникают дополнительные проблемы, связанные с присутствием злоумышленника в канале передачи ценной информации. Дистанционный мониторинг должен обеспечивать безопасную передачу как параметров соединения и текстов запросов, так и сведений о состоянии в открытой среде передачи, чтобы злоумышленник не смог получить доступ к удаленному SQL серверу и реконструировать структуру базу данных, а также перехватить ценные данные. Вместе с тем, от системы мониторинга требуется оперативное получение сведений для того, чтобы наиболее эффективно оценить работоспособность информационной среды и вовремя предотвратить возникающие проблемы.

Существует множество событий, которые могут снизить производительность SQL сервера [4-5]. Некоторые из них вызываются некорректной работой встроенных механизмов СУБД. Как известно, целостность хранимой информации обеспечивается механизмом блокировок СУБД, предотвращающим одновременное выполнение противоречивых операций над одними и теми же данными. Однако, долгое удержание запроса и чрезмерное количество блокировок может недопустимо снизить производительность СУБД [6]. Иными словами, требования к обеспечению целостности данных и высокой производительности противоречивы и согласовать их принципиально можно только поиском компромисса. Исходно, границы для поиска компромисса устанавливаются разработчиком СУБД на основании известных только ему условий и требований. Рабочие и, возможно, искусственно созданные злоумышленником запросы к БД могут эксплуатировать указанные границы компромиссного решения. В этих условиях мониторинг критических информационных процессов, имеющих отношение к блокировкам, становится актуальной проблемой. Поиск, разработка метода безопасного дистанционного мониторинга с глубоким и точным, но оперативным анализом блокирующих процессов является актуальной задачей.

Формулировка задачи

Поиск компромиссного решения сформулированных задач, очевидно, должен начаться с анализа приемлемых конкурентноспособных решений каждой из них по отдельности. Первой из них является обнаружение и анализ блокирующих процессов. Базовым способом решения является непрерывная, периодическая или выборочная оценка состояния сервера, а средством – использование специальных встроенных в SQL Server средств администрирования. Однако, базовые способы в той или иной степени снижают рабочую производительность сервера, а применение встроенных административных средств требует непосредственного доступа администратора к панели управления сервером. И то, и другое неприемлемо в контексте сформулированной выше цели данного исследования. Следовательно, единственным способом достичь цели является создание низкоуровневых точечных дистанционных командных воздействий на СУБД по извлечению необходимых и достаточных сведений о состоянии сервера с помощью специальных SQL-запросов.

Данные, содержащие сведения о состоянии SQL Server, можно получить из системных БД, обращение к которым осуществляется с помощью специальных запросов на языке SQL [7]. Среди множества набора средств, предназначенных для мониторинга состояния, стоит выделить (https://learn.microsoft.com/en-us/sql/relational-databases/performance/ performance-monitoring-and-tuning-tools):

- системные динамические административные представления (VIEW) [8];
- системные функции;
- системные хранимые процедуры.

Следует отметить, что основным критерием выбора объектов SQL Server является возможность быстро оценить состояние сервера с учетом полноты получаемых данных. В условиях дистанционного мониторинга следует также обеспечить безопасную передачу этих сведений по существующему каналу.

Анализ показал явное преимущество системных хранимых процедур. Они предоставляют запрашиваемые сведения одной командой, в которую на стороне сервера заложены необходимые операторы, недоступные для анализа в канале передачи данных. В этом состоит системный подход к достижению желаемого компромисса между требованиями обеспечения целостности данных и высокой производительности сервера. Этим достигается сокрытие содержание управляющих действий. Такие хранимые процедуры существуют, однако требуют существенной переработки или замены вследствие того, что были разработаны для решения универсальных задач мониторинга, или просто не отвечают текущим требованиям производительности.

Основой для рассмотрения и сравнительного анализа выбран ряд существующих хранимых процедур, выполняющих задачу мониторинга блокирующих процессов:

- sp_who (документированная);
- sp_who2 (недокументированная);
- sp_whoIsActive (http://whoisactive.com/).

Хранимые процедуры sp_who и sp_who2 основаны на устаревшем способе реализации с использованием системных

таблиц, разработчики от Microsoft рекомендуют использовать системные динамические представления. Однако, использование динамических представлений потребует создание конструкции с набором необходимых параметров в запросе хранимой процедуры.

Аналогичная функциональность с использованием системных динамических представлений присутствует в авторской хранимой процедуре sp_whoIsActive. Однако, запрашиваемые сведения для решения поставленной задачи сильно избыточны, что требует дополнительной фильтрации набора критических параметров.

Таким образом, следует создать новую хранимую процедуру, которая основана на рекомендациях от Microsoft с использованием системных динамических представлений и возвращает только необходимые сведения.

Второй задачей является создание или применение наиболее производительного и одновременно безопасного способа дистанционного информационного взаимодействия. Хотя нет оснований полностью отказываться от применения известных способов защиты информации в виде организации закрытого (шифрованного) канала, акцент в данной работе сделан на системных решениях, в частности, на использование высокоуровневых протоколов и многоуровневой клиент-серверной архитектуры. Ожидается, что высокоуровневый протокол обеспечит реализацию взаимодействия клиента с удалённым сервером с помощью высокоуровневых команд, а многоуровневая клиент-серверная система – опосредованное взаимодействие клиента с СУБД через промежуточный специализированный сервер приложений.

Обоснование структуры и состава параметров конструкции запроса

Случившаяся на сервере блокировка удерживает посылаемый запрос, поэтому задача отслеживания блокирующих процессов переходит в задачу оценивания производительности запроса. Для этого будет необходимо и достаточно определить время удержания запроса и реакцию ресурса SQL Server в виде затрат времени центрального процессора (ЦП) на исполнение запроса [9]. Дополнительно, для улучшения качества и оперативности оценки состояния SQL Server, в параметры состояния целесообразно включить тексты блокирующего и заблокированного запросов. Эти параметры позволят провести более точную оценку причин возникновения перегрузки из-за блокировок. Чтобы впоследствии устранить ожидающий запрос, снижающий производительность, следует указывать ID процесса. Логин пользователя позволит выявить подключение к серверу нежелательного пользователя [10].

В таблице 1 представлены необходимые и достаточные сведения о блокировках, отвечающие текущим требованиям производительности дистанционного мониторинга.

Поле *field* отображает соответствующее поле системного объекта СУБД, а поле *purpose* поясняет значение этого поля данных в процессе мониторинга. Так как административный запрос представляет собой хранимую процедуру, то следующим этапом является выбор необходимых системных динамических представлений, которые смогут обеспечить отображение предложенной структуры.

Таблица 1

Структура таблицы для мониторинга блокирующих запросов

№	field	data type	purpose
1	login	nvarchar	Логин пользователя
2	Blocking_ID	smallint	ID блокирующего процесса
3	Blocked_ID	smallint	ID заблокированного процесса
4	CPU	int	Время ЦП, затраченное на выполнение запроса
5	Wait_time_s	int	Время текущего ожидания заблокированного запроса (сек.)
6	DB	sysname	Имя базы данных
7	Status	nvarchar	Состояние процесса. Параметр необходим для фильтрации с использованием конструкции WHERE. Поиск запросов в ожидании.
8	Blocked_SQL_TEXT	nvarchar	Текст запроса в заблокированном процессе.
9	Blocking_SQL_TEXT	nvarchar	Текст запроса в блокирующем процессе.

Среди существующих динамических представлений, которые заменяют устаревшие системные таблицы, используемые в хранимых процедурах, описанных ранее, полноту возвращаемых данных может обеспечить системное динамическое представление: *sys.dm_exec_requests* (https://learn.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-exec-requests-transact-sql).

Данное системное представление обеспечивает необходимое количество связей с нужными полями данных из других системных представлений, а также обеспечивает фильтрацию для нахождения заблокированного запроса с помощью поля *status* и дополнительной конструкции WHERE.

Созданная конструкция в виде хранимой процедуры с именем *sp_FINDBLOCK* помещается в системную базу данных *master* (https://learn.microsoft.com/en-us/sql/t-sql/statements/create-procedure-transact-sql).

На рисунке 1 представлен результат выполнения данной процедуры в среде Microsoft SQL Server Management Studio в отсутствии блокирующих процессов.

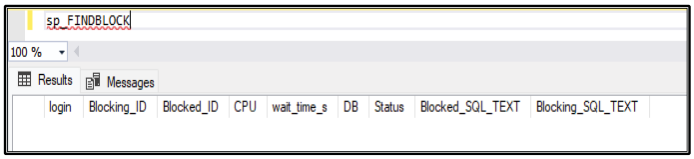


Рис. 1. Тестирование созданного запроса в отсутствии блокирующих процессов

Тестирование запроса в присутствии блокирующих процессов должно быть выполнено с участием специального сценария блокирования, который будет описан далее в соответствующем разделе.

Структура многоуровневой клиент-серверной информационной системы, используемой для тестирования средств удаленного мониторинга

В качестве инструмента для удаленного тестирования созданной инструкции используется авторское клиент-серверное приложение, подробно описанное в работе [11]. Сторона клиента выступает в качестве административной панели с возможностью использования запроса к SQL Server в виде команд на понятном языке пользователю (рис. 2).

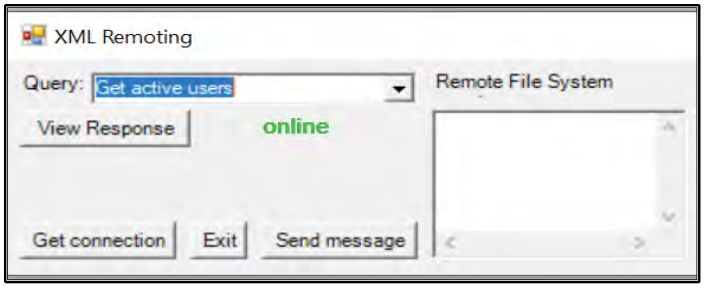


Рис. 2. Диалоговое окно клиентского приложения

Архитектура клиент-серверного приложения представляет собой четырехуровневую модель Web-доступа к удаленным данным (рис. 3), в которой по сети передаются не сами запросы в виде команд, а их идентификаторы, однозначно отображаемые в запросы на стороне сервера приложений с помощью высокоуровневого протокола.

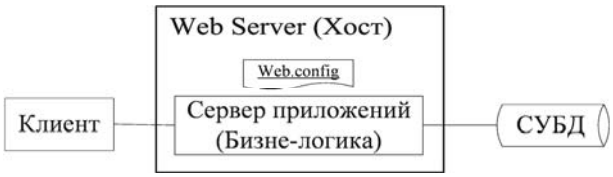


Рис. 3. Модель WEB-доступа к данным SQL Server

Инструкции SQL хранятся в XML-файле высокоуровневого протокола. Ниже приведена структура одного элемента протокола:

<XMLSQLqMaps>
<id>index</id>
<desc>description</desc>
<query>actual SQL query</query>
<rFile>filename</rFile>
</XMLSQLqMaps>.

Он содержит четыре параметра. Из них пользовательскими являются идентификатор *id*, словесное описание *desc* и имя файла *rFile*, в котором сохраняется полученный ответ. Идентификатор, как указано выше, однозначно отображается на стороне сервера приложений в SQL-запрос *query*, после чего посылается им на SQL Server по защищенному каналу.

За вход в систему и настройку ее параметров (строка соединения с SQL Server, его расположение в сети, параметры хостирования сервера приложений на Web-сервере и др.) отвечает конфигурационный файл *web.config*.

Предложенная авторская клиент-серверная система благодаря четырёхуровневой архитектуре и применению высокоуровневого протокола не использует передачу управляющих данных по открытому каналу, предотвращая тем самым возможности реконструировать базу данных или найти ее расположение.

Разработка тестовых сценариев блокировок, снижающих производительность SQL Server

Так как блокировки происходят при работе с большим объемом данных и большим количеством пользователей, то для повышения точности результатов тестирования выполнена имитация создания блокировок.

Первый способ принудительного создания блокировки – незавершенная транзакция (рис. 4). Задание инструкции *set_lock_timeout* не требуется, так как по умолчанию параметр принимает бесконечное значение. Пока данная транзакция не будет завершена, любые действия с указанной таблицей будут заблокированы.

```
use RemotingServerTest;
INSERT INTO USERS (Name) VALUES ('UserName');
BEGIN TRAN
UPDATE USERS SET Name='User_1' WHERE Name='UserName'
```

Рис. 4. Создание блокирующей инструкции

Второй способ заблокировать инструкцию – поставить запрос в ожидание с помощью конструкции *WAITFOR DELAY 'hh:mm:ss'*, т.е. выполнить принудительное удержание запроса (рис. 5).

```
use testBase;
BEGIN TRAN
UPDATE TEST set Test='Data_1' where TestID=1;
WAITFOR DELAY '00:20:10'
```

Рис. 5. Принудительное удержание запроса

Оценка эффективности разработанного метода

Ключевыми критериями эффективности использования конечных конструкций для дистанционного мониторинга являются: необходимое и достаточное количество сведений о состоянии (параметров состояния) и минимальное количество передаваемых сведений в самом запросе. Для созданных тестовых сценариев сравним возможности хранимых процедур sp_who2 и sp_WhoIsActive в обнаружении блокирующих процессов.

Sp_who2 отображает информацию обо всех процессах, происходящих на сервере, включая системные процессы, активируемые и регистрируемые базой данных *master* (поле *DBName*), как показано на рисунке 6. Информация о блокирующих процессах содержится в полях (столбцах таблицы) *BlkBy*, *STATUS* и *Command*. Потенциально наиболее информативным является содержание поля *Command*, однако в нем отображается только тип запроса (в нашем случае *SELECT*), а не сам текст запроса. Этих сведений недостаточно для того, чтобы провести полный анализ произошедшей блокировки. Таким образом, в sp_who2 отсутствует возможность просмотра заблокированного процесса, а также полного текста блокирующего запроса. Для выделения блокирующего процесса из всего множества процессов потребуется создание

временной таблицы [12], а это увеличит время обработки запроса СУБД и потребление оперативной памяти.


sp_who2

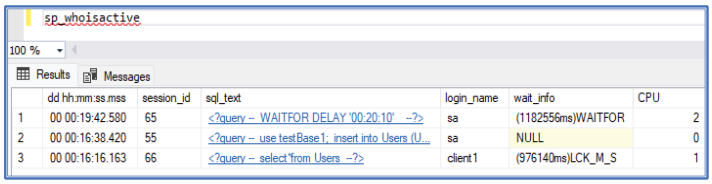
100 %

Results Messages

	SPID	Status	Login	HostName	BlkBy	DBName	Command	CPUTime	DiskIO	LastBatch	ProgramName
53	53	sleeping	sa	LAPTOP...		testBas...	AWAITING COMMAND	0	0	05/05 19:43:28	Microsoft S...
54	54	sleeping	sa	LAPTOP...		testBas...	AWAITING COMMAND	32	0	05/05 19:44:19	Microsoft S...
55	55	sleeping	sa			master	TASK MANAGER	0	0	05/05 19:49:06	
56	56	sleeping	sa	LAPTOP...		master	AWAITING COMMAND	0	0	05/04 15:30:04	Microsoft S...
57	57	SUSPENDED	sa	LAPTOP...	54	testBas...	SELECT	0	0	05/05 19:44:20	Microsoft S...

Рис. 6. Результат выполнения хранимой процедуры sp_who2

Сведения, предоставляемые sp_WhoIsActive без фильтрации по умолчанию также объемны, но включают в себя полный текст SQL-запроса (http://whoisactive.com/docs/19_whyblocked/), как показано на рисунке 7.


sp_whoisactive

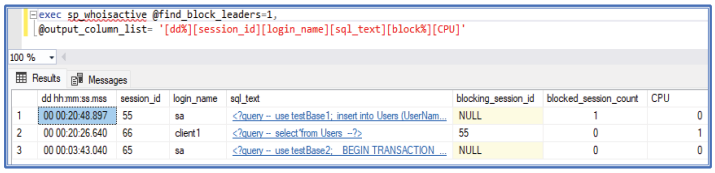
100 %

Results Messages

	dd hh:mm:ss.mss	session_id	sql_text	login_name	wait_info	CPU
1	00 00:19:42.580	65	<?query -- WAITFOR DELAY '00:20:10' --?>	sa	(1182556ms)WAITFOR	2
2	00 00:16:38.420	55	<?query -- use testBase1; insert into Users (U...	sa	NULL	0
3	00 00:16:16.163	66	<?query -- select*from Users --?>	client1	(976140ms)LCK_M_S	1

Рис. 7. Результат выполнения хранимой процедуры sp_WhoIsActive

Результаты исполнения хранимой процедуры sp_whoIsActive можно отфильтровать, используя параметры, представленные на рисунке 8.


exec sp_whoisactive @find_block_leaders=1,
@output_column_list= '[dd%][session_id][login_name][sql_text][block%][CPU]'

100 %

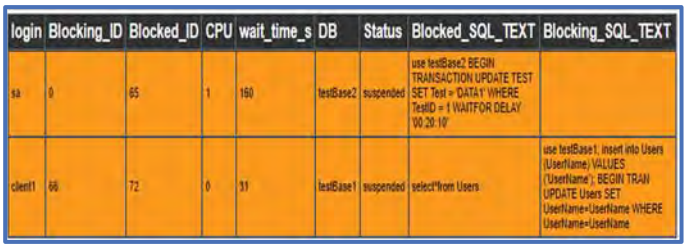
Results Messages

	dd hh:mm:ss.mss	session_id	login_name	sql_text	blocking_session_id	blocked_session_count	CPU
1	00 00:20:48.897	55	sa	<?query -- use testBase1; insert into Users (UserNam...	NULL	1	0
2	00 00:20:26.640	66	client1	<?query -- select*from Users --?>	55	0	1
3	00 00:03:43.040	65	sa	<?query -- use testBase2; BEGIN TRANSACTION ...	NULL	0	0

Рис. 8. Результат фильтрации sp_WhoIsActive

Однако, оптимальные комбинации параметров фильтрации определяются целью анализа и контекстом блокирующих процессов, а это затрудняет автоматизацию мониторинга. Оперативность и полнота мониторинга достигается не только запросом необходимого и достаточного количества сведений, но и компактностью самой конструкции запроса. Кроме того, видимость фильтрации столбцов в запросе (рис. 8) не отвечает требованиям безопасности в открытой среде передачи.

Преимущество новой созданной хранимой процедуры sp_FINDBLOCK перед конкурентноспособной sp_WhoIsActive – это сокрытие фильтрации сведений о состоянии в условиях дистанционной передачи. Она разделяет поток возвращаемой информации на потоки блокирующей и заблокированной инструкций, для цели более точного последующего анализа специалистом. Оперативность мониторинга достигается путем использования готовой отфильтрованной хранимой процедуры без необходимости дополнительной фильтрации запроса. Результат выполнения созданной хранимой процедуры в клиент-серверном приложении (рис. 9).



login	Blocking_ID	Blocked_ID	CPU	wait_time_s	DB	Status	Blocked_SQL_TEXT	Blocking_SQL_TEXT
sa	0	65	1	160	testBase2	suspended	use testBase2 BEGIN TRANSACTION UPDATE TEST SET Test = 'DATA1' WHERE TestID = 1 WAITFOR DELAY '00:20:10'	
client1	66	72	0	31	testBase1	suspended	select*from Users	use testBase1; insert into Users (UserName) VALUES ('UserName'); BEGIN TRAN UPDATE Users SET UserName=UserName WHERE UserName=UserName

Рис. 9. Результат выполнения sp_FINDBLOCK

Как видно, результат выполнения sp_FINDBLOCK отображается на колонки с префиксами «Blocking» и «Blocked», содержащие потоки блокирующей и заблокированной инструкций, что позволит специалисту выполнить их более глубокий анализ и оценку для более точного выявления причины появления блокировки.

Заключение

В статье выявлено, что все хранимые процедуры обладают способностью скрывать полный запрос и могут быть модифицированы с целью снижения риска раскрытия опасных сведений. Однако, рассмотренные конкурентно-способные хранимые процедуры не решают поставленную задачу нахождения компромисса между требованиями обеспечения целостности данных и высокой производительности сервера баз данных в условиях дистанционного мониторинга.

Выполнен сравнительный анализ известных конкурентноспособных хранимых процедур – *sp_who, sp_who2* и *sp_whoIsActive*. Показано, что процедуры *sp_who* и *sp_who2* нецелесообразно использовать в качестве основы для создания модифицированной хранимой процедуры из-за устаревшего и нерекомендуемого самим разработчиком способа реализации. Хранимая процедура *sp_whoIsActive* свободна от данного недостатка, но не обладает желаемым набором полезных параметров для достижения полноты мониторинга.

Высокая скрытность чувствительных сведений о структуре БД, достигаемая использованием хранимых процедур, может быть повышена путем их инкапсуляции в высокоуровневый протокол командного взаимодействия, реализуемого авторской клиент-серверной системой дистанционного мониторинга, с помощью которой выполнена детальная и успешная апробация разработанного инструментального средства в виде хранимой процедуры sp_FINDBLOCK.

Созданный административный запрос базируется на динамических административных представлениях, обеспечивая полноту мониторинга путём тщательного отбора признаков блокирующих процессов. Все параметры инкапсулированы в хранимую процедуру, позволяя скомпилировать созданную конструкцию только один раз с учетом сокрытия искомых сведений. Ключевой особенностью созданной хранимой процедуры является отображение блокирующей и заблокированной инструкции на языке SQL, что позволяет специалисту наиболее оперативно и точно проанализировать причину блокировки.

Литература

1. *Batra R.* A History of SQL and Relational Databases // Pro SQL Primer. Apress, Berkeley, CA, 2018, pp. 183-187. DOI: 10.1007/978-1-4842-3576-8_19

2. *Kate A.* et al. Conversion of natural language query to SQL query // 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2018, pp. 488-491. DOI: 10.1109/ICECA.2018.8474639

3. *Медведев Ю.С.* К вопросу об увеличении производительности базы данных Oracle // Экономическое развитие: состояние, проблемы, перспективы: Сборник статей Международной научно-практической конференции, Пенза, 28-29 июня 2018 года. Пенза: Автономная некоммерческая научно-образовательная организация «Приволжский Дом знаний», 2018. С. 71-75. EDN YMJLKP.

4. *Jung J., Hu H., Arulraj J., Kim T., Kang W.* APOLLO: Automatic Detection and Diagnosis of Performance Regressions in Database Systems (to appear) // Proceedings of the 46th International Conference on Very Large Data Bases (VLDB), Tokyo, Japan, Aug. 2020.

5. *Ilić M., Kopanja L., Zlatković D., Trajković M., Ćurguz D.* Microsoft SQL Server and Oracle: Comparative performance analysis // The 7th International conference Knowledge management and informatics. 2021, pp. 33-40.

6. *LaRock T., van de Laar E.* Wait Statistics Internals // Pro SQL Server 2022 Wait Statistics: A Practical Guide to Analyzing Performance in SQL Server and Azure SQL Database. Berkeley, CA : Apress, 2023. pp. 3-25.

7. *Zhao Z.* et al. T-SQL: A Lightweight Implementation to Enable Built-in Temporal Support in MVCC-Based RDBMSs // IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 1, pp. 1028-1042, 1 Jan. 2023, DOI: 10.1109/TKDE.2021.3081717.

8. *LaRock T., van de Laar E.* Querying SQL Server Wait Statistics // Pro SQL Server Wait Statistics: A Practical Guide to Analyzing Performance in SQL Server and Azure SQL Database. Berkeley, CA : Apress, 2023, pp. 27-63.

9. *Третьяков И.А., Кожекина Е.Н., Журавлев И.В.* Оптимизация SQL-запросов // Вестник Донецкого национального университета. Серия Г: Технические науки. № 2. 2021. С. 39-49. EDN RPSKQQ.

10. *Fleiner R., Hubert R., Bánáti A., Erdôdi L.* Security threats based on critical database system privileges // 2022 IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC), Reykjavík, Iceland, 2022, pp. 000117-000122, DOI: 10.1109/ICCC202255925.2022.9922750.

11. *Михайлов В.Ю., Мазепа Р.Б.* Практикум последовательности дисциплин в форме проектирования системы защищенного информационного взаимодействия в открытых сетях // Информационное противодействие угрозам терроризма. 2015. Т. 2, № 25. С. 161-169. EDN SYZGCX

12. *Вакульчик О.В.* Кибербезопасность функционирования информационно-управляющей системы с участием SQL server // Труды МАИ. 2022. № 127. DOI: 10.34759/trd-2022-127-13

REMOTE MONITORING OF EVENTS THAT CAUSE A DECREASE IN THE PERFORMANCE OF INFORMATION AND INFORMATION SEARCH SYSTEMS

VLADIMIR YU. MIKHAYLOV
Moscow, Russia

ROMAN B. MAZEPA
Moscow, Russia

OLGA V. VAKULCHIK
Moscow, Russia

KEYWORDS: *iremote status monitoring, remote administration, data integrity, SQL Server locking mechanism, SQL query.*

ABSTRACT

Introduction. The object of research is events that randomly or intentionally exploit the built-in locking mechanism of SQL Server, which significantly reduce the performance of information and information search engines. The subject of the research is the remote monitoring of blocking processes. **The purpose of the research** is to develop a secure remote monitoring method that can quickly and reliably detect events that cause the operation of the built-in locking mechanism of SQL Server. The use of monitoring is necessary to maintain the required level of system performance, as well as to promptly detect impacts on the information system that destroy the integrity of databases. This article is aimed at solving the main problems of remote monitoring: ensuring the completeness of the assessment of the condition and reliable, prompt transmission of the received information through an open channel. The key feature of the proposed remote monitoring method is the identification of SQL Server objects and SQL language constructs used to implement remote monitoring, taking into account the completeness of the information received. A high degree of security of the remote monitoring information environment is provided by the system means of implementing the author's client-server architecture. It is shown that the efficiency and security of the transfer of valuable control data is ensured by the use of stored procedures. Known stored procedures are redundant, so the authors needed to create their own monitoring tool, which is the main result of this article. **The results of the research** can be applied in the design of reliable and secure information and information search systems for various purposes.

REFERENCES

1. R. Batra, "A History of SQL and Relational Databases". *SQL Primer. Apress*, Berkeley, CA, 2018, pp. 183-187. DOI: 10.1007/978-1-4842-3576-8_19

2. A. Kate et al., "Conversion of natural language query to SQL query." 2018 Second International Conference on Electroni*cs, Communication and Aerospace Technology (ICECA)*. IEEE, 2018, pp. 488-491. DOI: 10.1109/ICECA.2018.8474639

3. Yu.S. Medvedev, "To the question about increasing productivity Oracle Database," *Ekonomicheskoe razvitie: sostoyanie, problemy, perspektivy,* Penza, 2018, pp. 71-75.

4. J. Jung, H. Hu, J. Arulraj, T. Kim, and W. Kang, “APOLLO: Automatic Detection and Diagnosis of Performance Regressions in Database Systems (to appear),” *Proceedings of the 46th International Conference on Very Large Data Bases (VLDB)*, Tokyo, Japan, Aug. 2020.

5. M. Ilic, L. Kopanja, D. Zlatkovic, M. Trajkovic, D. Curguz, "Microsoft SQL Server and Oracle: Comparative performance analysis," *The 7th International conference Knowledge management and informatics*. 2021, pp. 33-40.

6. T.LaRock, E. van de Laar, "Wait Statistics Internals," *Pro SQL Server 2022 Wait Statistics: A Practical Guide to Analyzing Performance in SQL Server and Azure SQL Database.* Berkeley, CA : Apress, 2023, pp. 3-25.

7. Z. Zhao et al., "T-SQL: A Lightweight Implementation to Enable Built-in Temporal Support in MVCC-Based RDBMSs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 1, pp. 1028-1042, 1 Jan. 2023, DOI: 10.1109/TKDE.2021.3081717.

8. T. LaRock, E. van de Laar, "Querying SQL Server Wait Statistics,” *Pro SQL Server Wait Statistics": A Practical Guide to Analyzing Performance in SQL Server and Azure SQL Database*. Berkeley, CA : Apress, 2023, pp. 27-63.

9. I.A. Tretiakov, E.N.Kozhekina, I.V. Zhuravlev, "Optimizing SQL-queries," *Bulletin of Donetsk National University. Series G: Technical Sciences*, no. 2, 2021, pp. 39-49.

10. R. Fleiner, R. Hubert, A. Banati and L. Erdodi, "Security threats based on critical database system privileges," *2022 IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC),* Reykjavik, Iceland, 2022, pp. 000117-000122, DOI: 10.1109/ICCC202255925.2022.9922750.

11. V.Yu. Mikhaylov, R.B. Mazepa, "Discipline sequence practice in the form of designing a secure information communication system in open networks", *Information counteraction to terrorist threats*, 2015, vol. 2, no. 25, pp. 161-169.

12. O.V. Vakulchik, “Cybersecurity of the functioning of the information management system with the SQL Server,” *Trudy MAI*, 2022, no. 127. DOI: 10.34759/trd2022-127-13

INFORMATION ABOUT AUTHORS:

[1] Professor, Dr. Sc. (Tech.), Moscow Aviation Institute (National Research University), Moscow, Russia, mihvj@yandex.ru

[2] Head the Department, Ph. D., Moscow Aviation Institute (National Research University), Moscow, Russia, mrb402@mai.ru

[3] Engineer, Moscow Aviation Institute (National Research University), Moscow, Russia, olga.vakulchik@mail.ru

For citation: Mikhaylov V.Yu., Mazepa R.B., Vakulchik O.V. Remote monitoring of events that cause a decrease in the performance of information and information search systems. H&ES Reserch. 2023. Vol. 15. No. 4. P. 68-74. doi: 10.36724/2409-5419-2023-15-4-68-74 (In Rus)