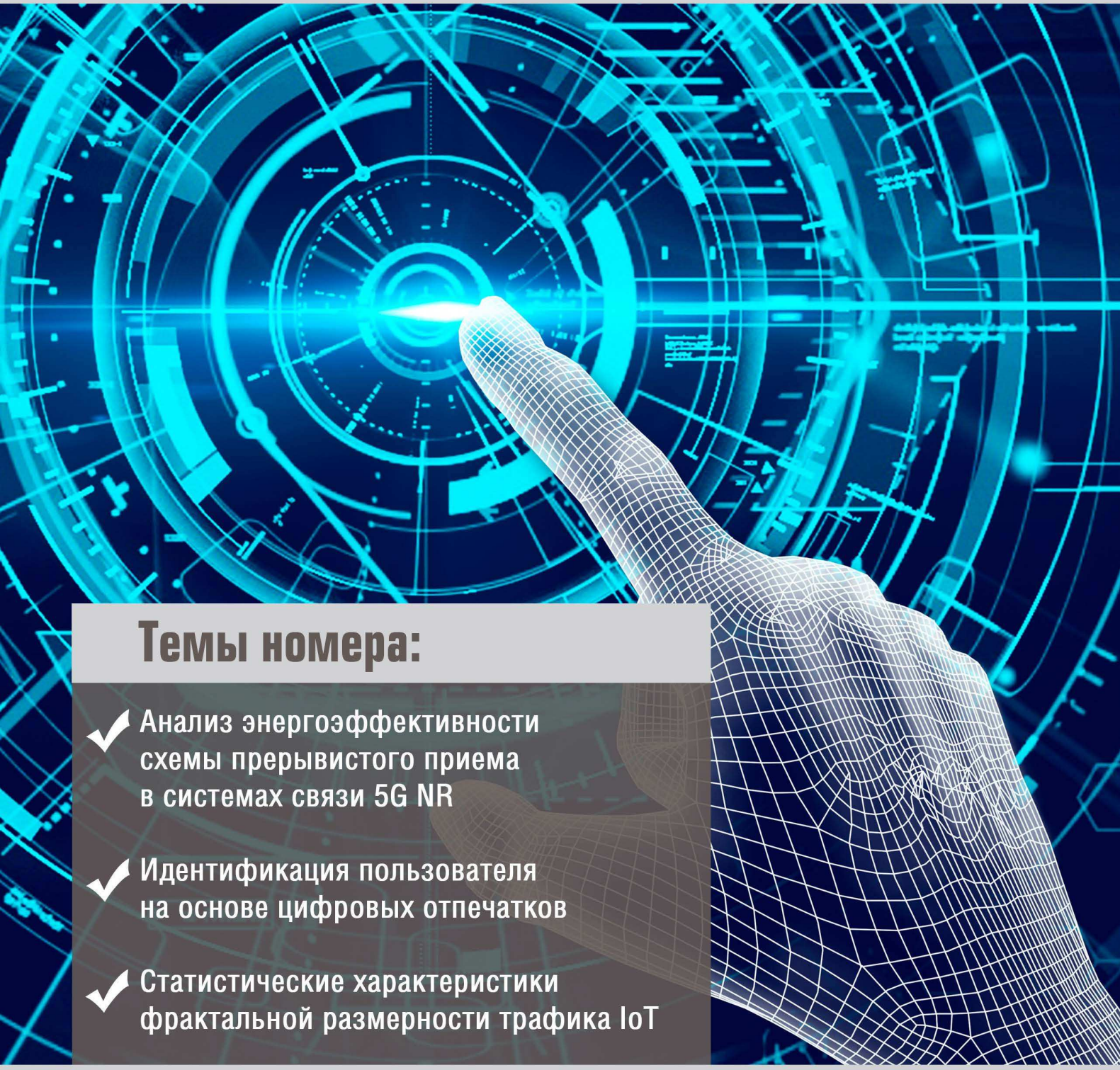
ISSN: 1813-324X (print)
2712-8830 (online)

ТРУДЫ УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Том 9. № 5
2023



Темы номера:

- Анализ энергоэффективности схемы прерывистого приема в системах связи 5G NR
- Идентификация пользователя на основе цифровых отпечатков
- Статистические характеристики фрактальной размерности трафика IoT

Vol. 9. Iss. 5
2023

PROCEEDINGS OF TELECOMMUNICATION UNIVERSITIES

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Научный журнал

ТРУДЫ
УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Том 9. № 5

Proceedings of Telecommunication Universities

Vol. 9. Iss. 5

Санкт-Петербург

2023

Труды учебных заведений связи. 2023. Т. 9. № 5. ISSN: 1813-324X (print), 2712-8830 (online)

Описание журнала

Научный журнал. Включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (распоряжение Минобрнауки России № 21-р от 12.02.2019), по специальностям (распоряжение № 33-р от 01.02.2022):

1.2.2. Математическое моделирование, численные методы и комплексы программ
2.2.6. Оптические и оптико-электронные приборы и комплексы
2.2.13. Радиотехника, в том числе системы и устройства телевидения
2.2.14. Антенны, СВЧ-устройства и их технологии
2.2.15. Системы, сети и устройства телекоммуникаций
2.2.16. Радиолокация и радионавигация
2.3.1. Системный анализ, управление и обработка информации
2.3.6. Методы и системы защиты информации, информационная безопасность

Выпускается с 1960 года. Выходит 6 раз в год. Издается на русском и английском языках.

Редакционный совет

Регистрационная информация

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций: ПИ № 77-77501 от 17.01.2020 г. (пред. рег. № 77-17986 от 07.04.2004 г.)

Подписной индекс в объединенном каталоге «ПРЕССА РОССИИ»: 59983

Размещение в РИНЦ (elibrary.ru) по договору: № 59-02/2013R от 20.02.2013

Контактная информация

Учредитель и издатель: Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Адрес учредителя: 191186, Санкт-Петербург, набережная реки Мойки, д. 61, литера А

Адрес редакции: 193232, Санкт-Петербург, пр. Большевиков, 22/1, к. 334/2

Тел.: +7 (812) 326-31-63, м. т. 2022, +79643759970

E-mail: tuzs@sut.ru

Web: http://tuzs.sut.ru

ВК: http://vk.com/spbtuzs

Proceedings of Telecommun. Univ. 2023. Vol. 9. Iss. 5. ISSN: 1813-324X (print), 2712-8830 (online)

Description

Scientific journal. The journal is included in the List of reviewed scientific publications, in which the main scientific results of dissertations for the degree of candidate of science and for the degree of doctor of science should be published (order of the Ministry of Education and Science of Russia No 21-r of 12 February 2019) in the field of (order of the Ministry of Education and Science of Russia No 33-r of 01 February 2022):

1.2.2. Mathematical modeling, numerical methods and complexes of programs
2.2.6. Optical and optoelectronic devices and complexes
2.2.13. Radio engineering, including television systems and devices
2.2.14. Antennas, microwave devices and its technologies
2.2.15. Systems, networks and telecommunication devices
2.2.16. Radiolocation and radio navigation
2.3.1. System analysis, management and information processing
2.3.6. Methods and systems of information security, cybersecurity

Since 1960. Published 6 times per year. Published in Russian and English.

Editorial Board

Registration Information

Registered by Federal Service for Supervision of Communications, Information Technology and Mass Media on 17.01.2020: PI No. 77-77501 (prev. reg. on 04.07.2004: No. 77-17986)

Subscription index for joint catalog «PRESSA ROSSII»: 59983

Accommodation in RINC (elibrary.ru) by agreement on 20.02.2013: No. 59-02/2013R

Contact Information

Publisher: Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruevich Saint-Petersburg State University of Telecommunications» (SPbSUT)

Publisher address: 191186, Saint Petersburg, Moika river embankment, 61-A

Post address: 193232, Saint Petersburg, Prospekt Bolshevikov, 22/1

Phone: +7 (812) 326-31-63, local 2022, +79643759970

E-mail: tuzs@sut.ru

Web: http://tuzs.sut.ru

СОДЕРЖАНИЕ

CONTENTS

ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ МОЛОДЫХ УЧЕНЫХ


ЭЛЕКТРОНИКА, ФОТОНИКА,
ПРИБОРОСТРОЕНИЕ И СВЯЗЬ
2.2.6 – Оптические и оптико-электронные приборы и комплексы
2.2.13 – Радиотехника, в том числе системы и устройства телевидения
2.2.14 – Антенны, СВЧ-устройства и их технологии
2.2.15 – Системы, сети и устройства телекоммуникаций
2.2.16 – Радиолокация и радионавигация

Научная статья
УДК 621.38
DOI:10.31854/1813-324X-2023-9-5-6-15

Последовательные аналого-цифровые КАМ-модемы на базе комплексных полосовых фильтров с НЧ-прототипами Баттерворта

Юрий Александрович Гребенко, GrebenkoYA@mpei.ru
Роман Игоревич Поляк ✉, poliakri@mpei.ru
Лин Пайнг Ян, hlainff@gmail.com

Национальный исследовательский университет «МЭИ»,
Москва, 111250, Российская Федерация

Аннотация: *Статья посвящена разработке и реализации алгоритмов квадратурной амплитудной модуляции на базе аналогового комплексного полосового фильтра с НЧ-прототипами Баттерворта. Использование комплексных полосовых фильтров Баттерворта позволяет получить несущие сигналы с практическими неперекрывающимися спектрами. При этом аналоговый комплексный фильтр Баттерворта имеет нелинейную фазочастотную характеристику и, соответственно, бесконечную и несимметричную импульсную характеристику. В статье проведен обзор разработки и реализации последовательного КАМ-модема на базе аналогового комплексного полосового фильтра Баттерворта и обратного цифрового КИХ-фильтра. Рассматривается процедура получения выражения импульсной характеристики аналоговых комплексных полосовых фильтров Баттерворта. С помощью схемотехнического моделирования определяются амплитудно-частотная и импульсная характеристики таких фильтров. Предложена структурная схема последовательной системы передачи данных с квадратурной амплитудной модуляцией на базе аналогового комплексного полосового фильтра Баттерворта и обратного цифрового КИХ-фильтра. Приведены результаты ее схемотехнического моделирования в среде Micro-Cap.*

Ключевые слова: *аналоговый комплексный полосовой фильтр, амплитудно-частотная характеристика, бесконечная импульсная характеристика, фазочастотная характеристика, квадратурная амплитудная модуляция, последовательный КАМ-модем*

Ссылка для цитирования: Гребенко Ю.А., Поляк Р.И., Ян Л.П. Последовательные аналого-цифровые КАМ-модемы на базе комплексных полосовых фильтров с НЧ-прототипами Баттерворта // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 6–15. DOI:10.31854/1813-324X-2023-9-5-6-15

Serial Analog-Digital QAM Modems Based on Complex Band-Pass Filters with LF Butterworth Prototypes

Yury Grebenko, GrebenkoYA@mpei.ru
Roman Polyak ✉**,** poliakri@mpei.ru
Yan Lin P., hlainff@gmail.com

National Research University "MPEI",
Moscow, 111250, Russian Federation

Abstract: The article is devoted to the development and implementation of quadrature amplitude modulation algorithms based on an analog complex band-pass filter with Butterworth low-frequency prototypes. The use of complex Butterworth band pass filters makes it possible to obtain carrier signals with practically non-overlapping

spectra. But the analog complex Butterworth filter has a non-linear phase-frequency response and, accordingly, the impulse responses of which are infinite and asymmetric. An overview of the development and implementation of a serial QAM modem based on an analog Butterworth complex band-pass filter and an inverse digital FIR filter. The procedure for obtaining the expression for the impulse response of analog complex band-pass Butterworth filters is considered. With the help of circuit simulation, the frequency response and impulse response of such a filter are determined. The choice of the center frequency is carried out by setting two coefficients. A block diagram of a serial data transmission system with quadrature amplitude modulation based on an analog complex Butterworth bandpass filter and an inverse digital FIR filter is proposed. The results of its circuit simulation in the Micro-Cap environment are presented.

Keywords: *analog complex bandpass filter, amplitude-frequency response, infinite impulse response, phase-frequency response, quadrature amplitude modulation, serial QAM modem*

For citation: Grebenko Y., Polyak R., Yan L.P. Serial Analog-Digital QAM Modems Based on Complex Band-Pass Filters with LF Butterworth Prototypes. *Proceedings of Telecommun. Univ.* 2023;9(5):6–15. DOI:10.31854/1813-324X-2023-9-5-6-15

Введение

Исследования различных телекоммуникационных технологий передачи сигналов показали, а практика подтвердила, что современным требованиям наилучшим образом отвечают системы связи, использующие для передачи сигналов множество ортогональных гармонических сигналов (переносчиков), одновременно и независимо модулируемых передаваемыми информационными сигналами. Ранее [1–3] были предложены способы реализации комплексных полосовых фильтров (ПФ), используемых в различных системах передачи сигналов. Таким образом, несущие сигналы с шириной спектров, ограниченной частотным диапазоном канала связи, используются при построении таких систем с полосно-ограниченным каналом связи [4].

Для реализации алгоритмов модуляции и демодуляции в этом случае применяют фильтровые методы. Известны фильтровые системы передачи данных на базе вещественных фильтров [4]. Систему передачи данных с квадратурной амплитудной модуляцией (КАМ) можно создать на базе аналоговых комплексных ПФ Баттерворта, позволяющих получить ортогональные сигналы. Если в модуляторе использовать несимметричную импульсную характеристику фильтра Баттерворта в качестве несущего сигнала, то в части демодулятора необходимо использовать согласованный цифровой комплексный КИХ-фильтр с усеченной импульсной характеристикой.

Достаточно часто применяется последовательная структурная схема системы передачи данных, состоящая из вещественных аналоговых комплексных ПФ Баттерворта. При последовательном способе передачи используется одноканальный модем с одним сигналом-переносчиком в общем случае конечной длительности, который одновременно и независимо модулируется в тактовые моменты передаваемым информационным сигналом. В статье предлагается использовать вариант аналогового комплексного ПФ Баттерворта в последовательной структурной схеме системы КАМ-модема.

Разработка аналогового комплексного полосового фильтра Баттерворта по координатам полюсов НЧ-прототипа

Рассмотрим НЧ-прототипы, описываемые только полюсами. В этом случае передаточная функция представляется в виде произведения сомножителей 1-го порядка $Ti(s)$ и может быть записана в виде следующего выражения [5]:

$$Ti(s) = \frac{k}{s + a_i + jb_i},$$

где k – коэффициент усиления.

Передаточная функция НЧ-прототипа Баттерворта 3-го порядка имеет вид:

$$T(s) = T1(s) \times T2(s) \times T3(s),$$

$$T(s) = \frac{1}{s+1} \times \frac{1}{s + 0.5 + j0.866} \times \frac{1}{s + 0.5 - j0.866},$$

где $a_1 = 1$; $b_1 = 0$; $a_2 = a_3 = 0.5$; $b_2 = 0.866$; $b_3 = -0.866$.

Передаточную функцию, соответствующую частотной характеристике, смещенной вправо на величину Ω_0, можно представить в виде выражения (1).

$$\bar{T}_1(s) = \frac{k}{s - j\Omega_0 + a_i + jb_i} = \frac{k}{s + a_i + j(b_i - \Omega_0)} = \frac{k\frac{1}{s+a_i}}{1 + j(b_i - \Omega_0)\frac{1}{s+a_i}} = \frac{\frac{k}{a_i}\frac{a_i}{s+a_i}}{1 + j\frac{(b_i-\Omega_0)}{a_i}\frac{a_i}{s+a_i}}. \quad (1)$$

Преобразованной передаточной функции $\bar{T}_1(s)$ можно поставить в соответствие структурную схему комплексного звена, показанную на рисунке 1 [5].

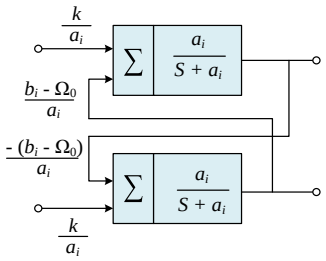


Рис. 1. Структурная схема НЧ-прототипа комплексного звена

Fig. 1. Block Diagram of the LF Prototype of the Complex Link

Структурную схему, показанную на рисунке 1, будем называть структурированным НЧ-прототипом комплексного звена. Для перехода к схеме, обеспечивающей заданные параметры комплексного ПФ, а именно полосу пропускания Δω и центральную частоту ω_0, надо выполнить замену переменной:

$$S = \frac{P}{\omega_п},$$

где $\omega_п = \frac{\Delta\omega}{2}$; при этом значение нормированного смещения соответствует выражению:

$$\Omega_0 = \frac{\Delta\omega}{\omega_п} = \frac{2\omega_0}{\Delta\omega} = \frac{2f_0}{\Delta f}.$$

На рисунке 2 приведена структурная схема суммирующего фильтра нижних частот (ФНЧ) 1-го порядка, которую можно реализовать с использованием программы схемотехнического моделирования Micro-Cap, где $f_{0i} = \frac{a_i \Delta f}{2}$.

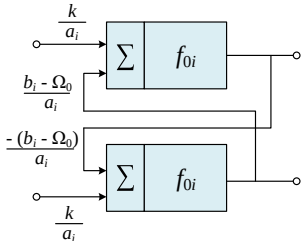


Рис. 2. Структурная схема комплексного звена

Fig. 2. The Block Diagram of the Complex Link

На рисунке 3 показана модель комплексного звена в программе Micro-Cap, где $\frac{k}{a_i} = \frac{R}{R_{ik}}$; $\frac{b_i - \Omega_0}{a_i} = \frac{R}{R_{iP}}$.

Разработанная в данной статье методика синтеза чаще всего применяется на низких частотах, поэтому для примера были выбраны центральная частота f_0 = 5 кГц и полоса пропускания Δf = 5 кГц в качестве параметров комплексного ПФ по аналогии с предыдущими работами авторов [6, 7].

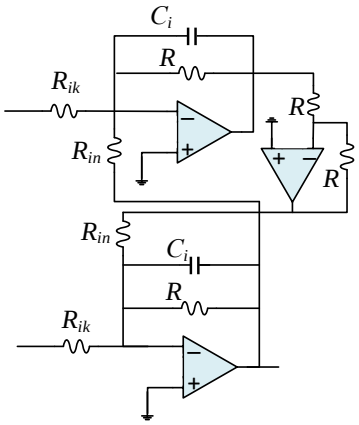


Рис. 3. Модель комплексного звена

Fig. 3. The Model of the Complex Link

Определяем нормированное смещение:

$$\Omega_0 = \frac{2f_0}{\Delta f} = \frac{2 * 5000}{5000} = 2.$$

Для 1-го звена ($a_1 = 1$; $b_1 = 0$), находим:

$$f_{01} = \frac{a_1 \Delta f}{2} = \frac{1 * 5000}{2} = 2500;$$

$$\frac{k}{a_1} = \frac{1}{1} = 1, \quad \frac{b_1 - \Omega_0}{a_1} = \frac{0 - 2}{1} = -2.$$

Структурная схема 1-го звена показана на рисунке 4.

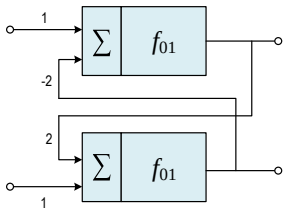


Рис. 4. Структурная схема 1-го звена

Fig. 4. Block Diagram of the 1st Link

Для 2-го звена ($a_2 = 0.5$; $b_2 = -0.866$) находим:

$$f_{02} = \frac{a_2 \Delta f}{2} = \frac{0.5 * 5000}{2} = 1250 \text{ Гц},$$

$$\frac{k}{a_2} = \frac{1}{0.5} = 2, \quad \frac{b_2 - \Omega_0}{a_2} = \frac{-0.866 - 2}{0.5} = -5.732.$$

Структурная схема 2-го звена показана на рисунке 5.

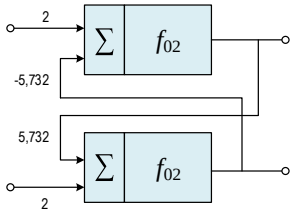


Рис. 5. Структурная схема 2-го звена

Fig. 5. Block Diagram of the 2nd Link

Для 3-го звена ($a_3 = 0.5$; $b_3 = 0.866$) находим:

$$f_{03} = \frac{a_3 \Delta f}{2} = \frac{0.5 * 5000}{2} = 1250 \text{ Гц},$$

$$\frac{k}{a_3} = \frac{1}{0.5} = 2, \quad \frac{b_3 - \Omega_0}{a_3} = \frac{0.866 - 2}{0.5} = -2.26.$$

Структурная схема 3-го звена показана на рисунке 6.

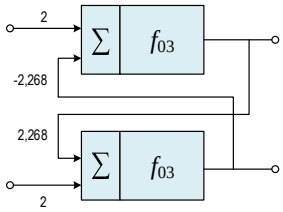


Рис. 6. Структурная схема 3-го звена

Fig. 6. Block Diagram of the 3rd Link

Полная структурная схема рассчитанного комплексного звена фильтра показана на рисунке 7. На рисунке 8 представлена принципиальная схема комплексного ПФ Баттерворта 3-го порядка в программе Micro-Cap (здесь и далее – не строго по ЕСКД).

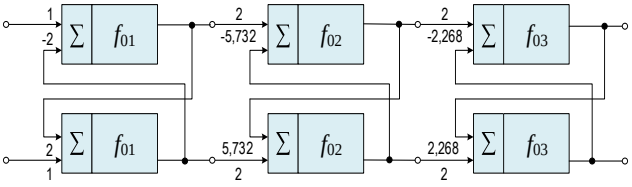


Рис. 7. Структурная схема комплексного полосового фильтра с НЧ-прототипом Баттерворта 3-го порядка

Fig. 7. Block Diagram of a Complex Bandpass Filter with a Low-Frequency Butterworth Prototype of the 3rd Order

Рассчитанная амплитудно-частотная характеристика (АЧХ) комплексного ПФ Баттерворта 3-го порядка показана на рисунке 9.

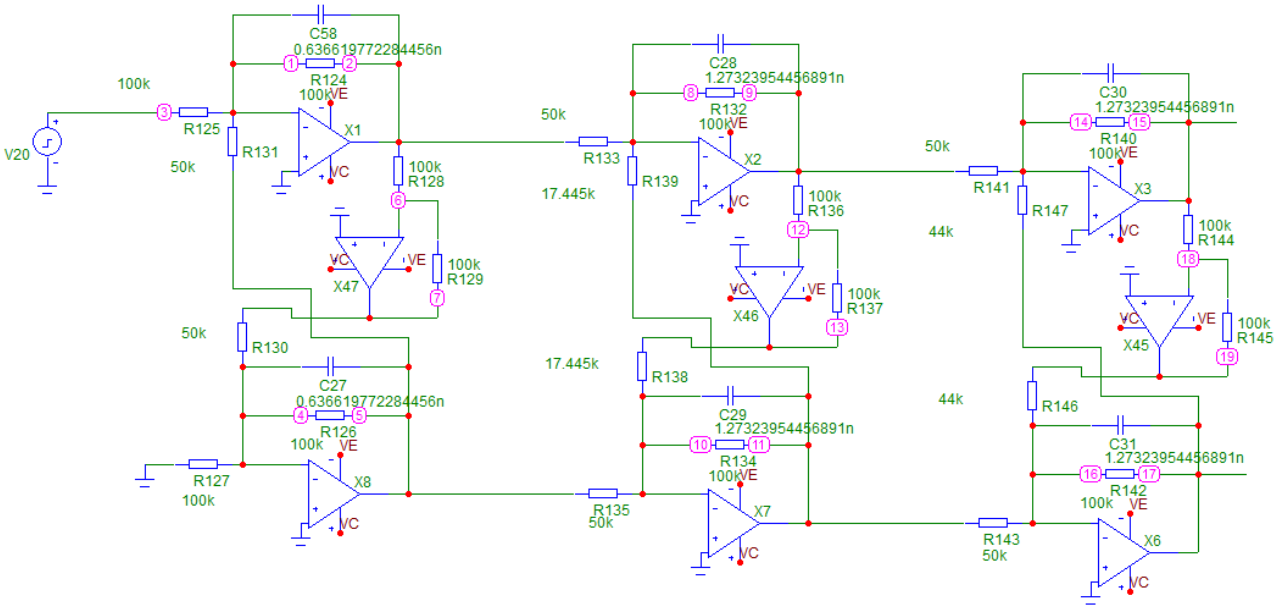


Рис. 8. Принципиальная схема комплексного полосового фильтра Баттерворта 3-го порядка с центральной частотой 5000 Гц

Fig. 8. Circuit Diagram of a Third-Order Butterworth Complex Bandpass Filter with a Center Frequency of 5000 Hz

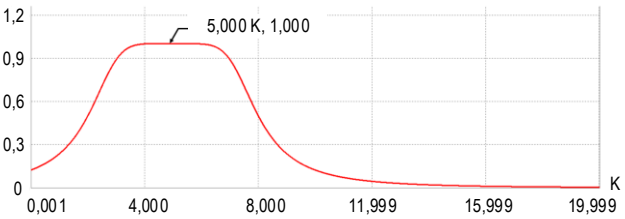


Рис. 9. АЧХ канала передачи с центральной частотой 5000 Гц

Fig. 9. Frequency Response of the Transmission Channel with a Central Frequency of 5000 Hz

Разработка цифрового комплексного полосового обратного КИХ-фильтра Баттерворта

Зададим следующие параметры фильтра: П = = 2500 Гц. ФНЧ будем проектировать на идентичных звеньях. В качестве передаточной функции НЧ-прототипа выступает функция Баттерворта 4-го порядка:

$$T_{\text{нч}} = \frac{1}{s^3 + 2s^2 + 2s + 1}.$$

Далее необходимо рассчитать параметры структурной схемы НЧ-прототипа на базе идентичных звеньев 1-го порядка. В качестве звена 1-го порядка выберем звено с передаточной функцией [8–10]:

$$K_{\text{звена}}(s) = \frac{1}{s + 1},$$

тогда

$$s(K) = \frac{(1 - K)}{K}.$$

Принципиальная схема идентичного звена показана на рисунке 10.

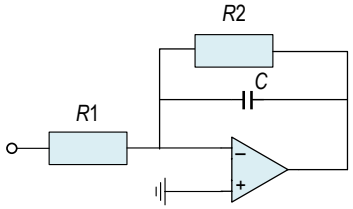


Рис. 10. Принципиальная схема идентичного звена

Fig. 10. Circuit Diagram of an Identical Link

Подставим передаточную функцию базового звена в передаточные функции блоков НЧ-прототипа (2).

Построим блок НЧ-прототипа по канонической структурной схеме. Структурные схемы, соответствующие передаточным функциям блоков, представлены на рисунке 11. Общая структурная схема разработанного фильтра строится путем последовательного соединения принципиальных схем 2-х блоков. На рисунке 12 представлена принципиальная схема ФНЧ Баттерворта 4-го порядка.

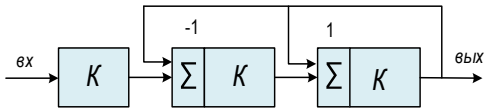


Рис. 11. Структурная схема НЧ-прототипа Баттерворта 4-го порядка

Fig. 11. Block Diagram of the Low-Frequency Prototype of Butterworth 4th Order

$$T_{\text{нч}} = \frac{1}{((1-K)/K)^3 + 2((1-K)/K + 2((1-K)/K) + 1} = \frac{K^3}{1-K+K^2}. \quad (2)$$

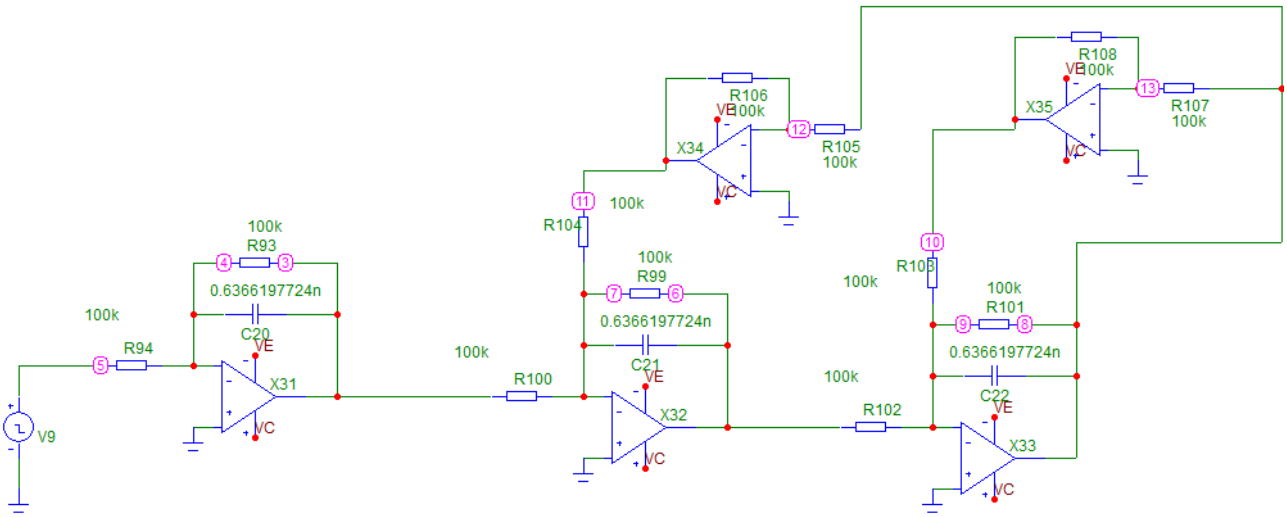


Рис. 12. Принципиальная схема ФНЧ Баттерворта 4-го порядка

Fig. 12. The 4th Order Butterworth LPF Circuit Diagram

Путем моделирования в среде Micro-Cap находим АЧХ (рисунок 13) и импульсную характеристику (рисунок 14) рассчитанного ФНЧ. Проведем дискретизацию полученной импульсной характеристики (рисунок 15), частота дискретизации равна $f_д$ = 80 кГц.

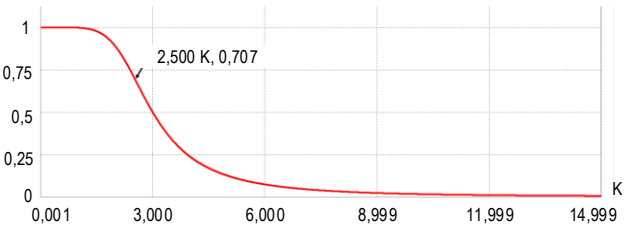


Рис. 13. АЧХ ФНЧ Баттерворта 4-го порядка (2.5 кГц)

Fig. 13. Frequency Response of the 4th Order Butterworth LF (2.5 k Hz)

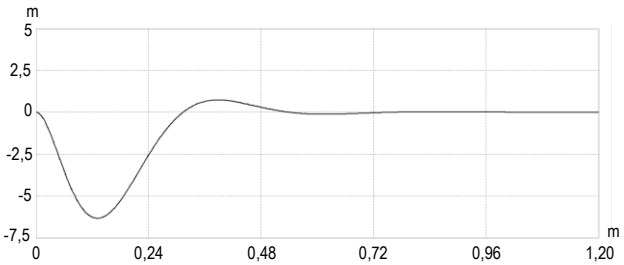


Рис. 14. Импульсная характеристика аналогового ФНЧ

Fig. 14. Impulse Response of the Analog LPF

Ограничим импульсную характеристику 64-м отсчетом. В этом случае импульсная характеристика будет представлена дискретной последовательностью $h(n) = \{h_1, h_2, h_3 \ldots h_{64}\}$.

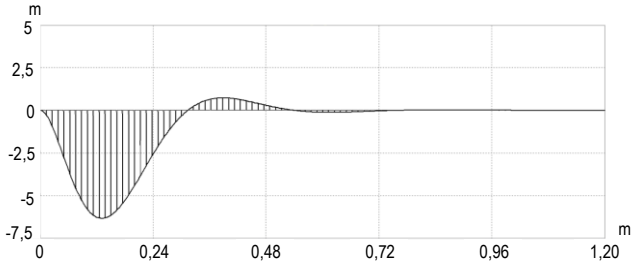


Рис. 15. Отсчеты импульсной характеристики ($f_д$ = 80 кГц)

Fig. 15. Impulse Response Readings (d = 80 k Hz)

Импульсная характеристика линеаризующего КИХ-фильтра имеет обратный порядок следования отсчетов $h(n) = \{h_{64}, h_{15}, h_{14} \ldots h_1\}$ [6, 7] (3). Таким образом передаточная функция обратного цифрового КИХ-фильтра Баттерворта с нерекурсивной формой представлена выражением (4). Структурная схема обратного цифрового ФНЧ КИХ-фильтра Баттерворта показана на рисунке 16. АЧХ такого фильтра практически совпадает с АЧХ аналогового ФНЧ (рисунок 17).

$$h(n) = \{0.013, 0.009, 0.003, -0.005, -0.014, -0.024, -0.036, -0.048, -0.061, -0.075, -0.088, -0.100, -0.111, -0.119, -0.124, -0.125, -0.122, -0.112, -0.096, -0.071, -0.039, 0.003, 0.053, 0.111, 0.177, 0.249, 0.326, 0.405, 0.484, 0.559, 0.626, 0.680, 0.716, 0.727, 0.711, 0.658, 0.564, 0.424, 0.231, -0.030, -0.339, -0.707, -1.131, -1.609, -2.133, -2.693, -3.276, -3.867, -4.445, -4.989, -5.474, -5.875, -6.166, -6.321, -6.323, -6.155, -5.803, -5.266, -4.557, -3.707, -2.763, -1.801, -0.924, -0.272\}. \quad (3)$$

$$T(z) = 0.013 + 0.009z^{-1} + 0.003z^{-2} + \ldots + -0.924z^{-63} - 0.272z^{-64}. \quad (4)$$

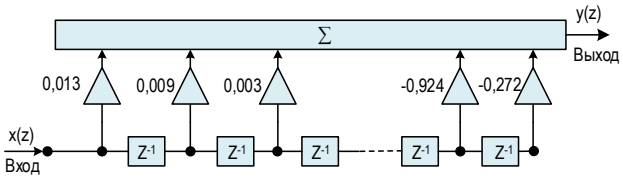


Рис. 16. Структурная схема обратного цифрового ФНЧ КИХ-фильтра Баттерворта

Fig. 16. Block Diagram of the Reverse Digital Low-Pass Filter Butterworth

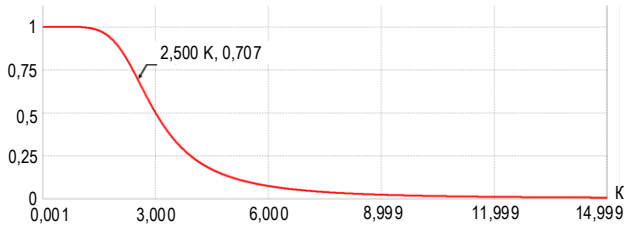


Рис. 17. АЧХ аналогового ФНЧ Баттерворта 4-го порядка с полосой 2.5 кГц и АЧХ КИХ-фильтра

Fig. 17. The Frequency Response of the 4th-Order Analog Low-Frequency Butterworth with a 2.5 kHz Band and the Frequency Response of the FIR Filter

Чтобы получить структурную схему комплексного цифрового ПФ, необходимо блоки задержек в структурной схеме ФНЧ заменить на комплексные задержки (рисунок 18) [8–10].

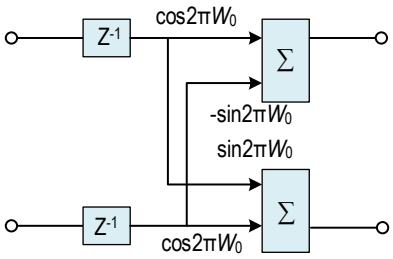


Рис. 18. Структурная схема комплексной задержки

Fig. 18. Block Diagram of Complex Delay

Пусть частотная характеристика смещается вправо на 5 кГц, тогда:

$$w_0 = \frac{f_c}{f_g} = 0.0625,\ \phi_0 = 2\pi w_0 = \frac{\pi}{2},\ e^{j\phi_0} = j.$$

Модель обратного цифрового комплексного полосового КИХ-фильтра Баттерворта показана на рисунке 19. Рассчитанная АЧХ такого цифрового комплексного полосового КИХ-фильтра Баттерворта показана на рисунке 20.

Можно констатировать, что АЧХ обратного цифрового КИХ-фильтра совпадает с АЧХ аналогового фильтра. Фазо-частотная характеристика (ФЧХ) и неравномерность зависимости группового времени запаздывания (ГВЗ) от частоты в полосе пропускания для аналогового комплексного фильтра Баттерворта приведены на рисунке 21. Оценим эффективность линеаризации при последовательном соединении обратного КИХ-фильтра. ФЧХ и ГВЗ для их последовательного соединения (рисунок 2).

Можно отметить, что ФЧХ стала более линейной, а неравномерность зависимости ГВЗ от частоты в полосе пропускания стала меньше.

Модем на основе аналого-цифровых комплексных полосовых фильтров

Следующим шагом является разработка модема на базе аналогового комплексного фильтра Баттерворта и обратного цифрового КИХ-фильтра. Его структурная схема, где $\delta(t - mT_{\text{ТАКТ}})$ – единичный импульс, m – номер тактового интервала, $T_{\text{ТАКТ}}$ – тактовый интервал поступления передаваемых символов, K – масштабный коэффициент амплитудной модуляции, показана на рисунке 23.

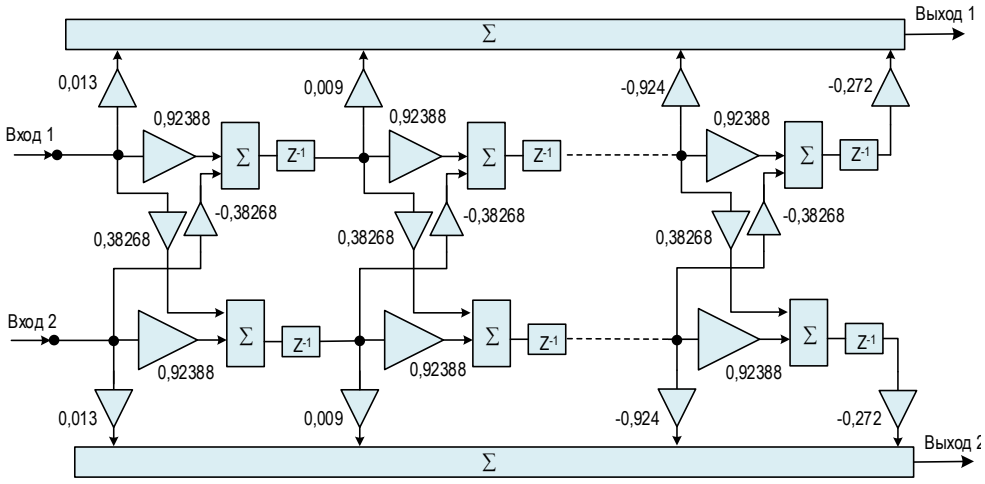


Рис. 19. Структурная схема обратного (согласованного) цифрового комплексного полосового КИХ-фильтра Баттерворта

Fig. 19. Block Diagram of the Reverse (Matched) Digital Complex Butterworth Bandpass Filter

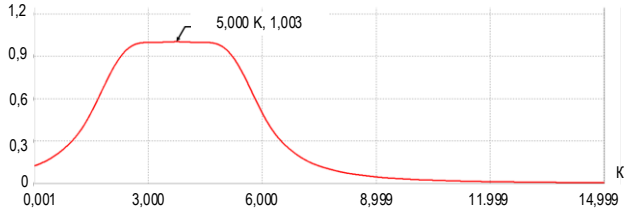
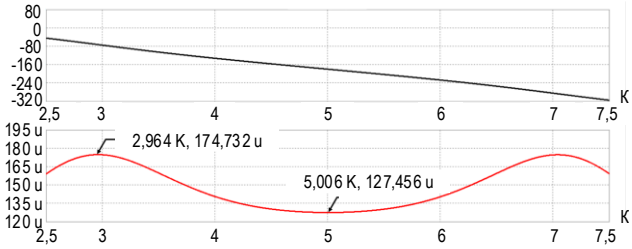


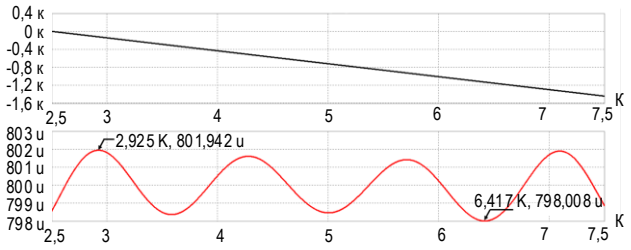
Рис. 20. АЧХ обратного (согласованного) цифрового комплексного полосового КИХ-фильтра Баттерворта

Fig. 20. Frequency Response of the Reverse (Matched) Digital Complex Butterworth Bandpass Filter

Модулятор состоит из аналогового комплексного ПФ Баттерворта, устройства взвешивания и выходного сумматора. При подаче единичного импульса на вещественный вход, на выходах канального фильтра возникнут сигналы, совпадающие с вещественной и мнимой частями импульсной характеристики комплексного фильтра. Чтобы сформировать КАМ-сигналы, необходимо умножить вещественную и мнимую части выходного сигнала комплексного фильтра на соответствующие масштабные коэффициенты.



a)



b)

Рис. 21. ФЧХ и зависимость ГВЗ от частоты в полосе пропускания: a) фильтра Баттерворта; b) фильтра Баттерворта и КИХ-фильтра

Fig. 21. The Frequency Response and the Dependence of the Group Delay Time on the Frequency in the Bandwidth: a) Butterworth Filter; b) Filter Butterworth and FIR-Filter

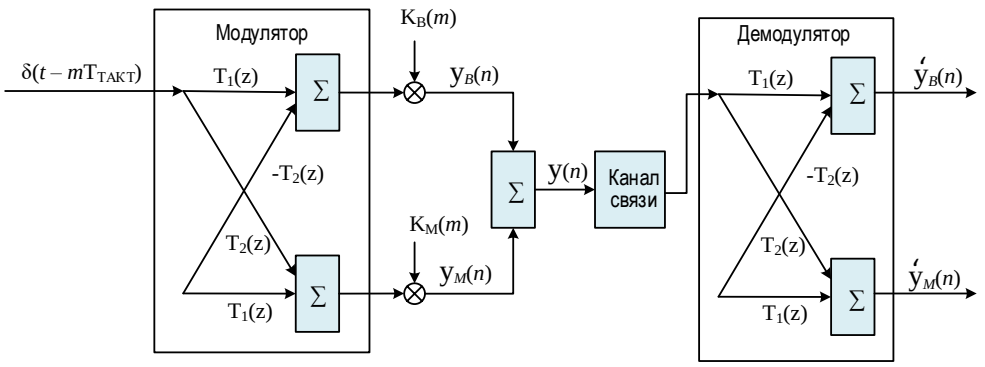


Рис. 23. Общая структурная схема системы передачи данных

Fig. 23. General Block Diagram of the Data Transmission System

В модуляторе используется несимметричная импульсная характеристика фильтра Баттерворта в качестве несущего сигнала, при этом необходимо в демодуляторе использовать согласованный цифровой комплексный КИХ-фильтр с усеченной импульсной характеристикой. В канал связи поступает общий суммарный сигнал. После прохождения через канал связи этот сигнал поступает на канальный фильтр демодулятора. Мы получаем по два сигнала на выходах канального обратного цифрового КИХ-фильтра демодулятора.

Сигнальное созвездие основы передачи КАМ

В модуляторе амплитуда ортогональных сигналов с выхода комплексного фильтра изменяется в соответствии с передаваемой информацией. Амплитуды ортогональных сигналов определяются с помощью сигнального созвездия для КАМ-16 (рисунок 24); взвешивающие коэффициенты $K_{\text{В}}$ и $K_{\text{М}}$ для каждой комбинации четырех бит можно записать в виде (5).

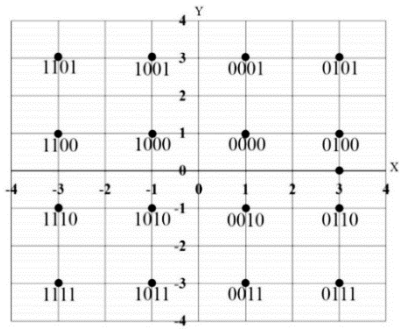


Рис. 24. Сигнальное созвездие КАМ-16

Fig. 24. Signal Constellation QAM-16

Формирование сигналов в модуляторе

Рассмотрим пример формирования сигналов в модуляторе при передаче комбинации из двенадцати битов 0001 0010 1100. В этом случае взвешивающие коэффициенты зависят от номера тактового интервала (m). В таблице 1 показан пример определения взвешивающих коэффициентов модуляции в соответствии с номером тактового интервала. Выберем тактовый интервал между входными импульсами, равным $T_{\text{ТАКТ}} = 1.5$ мс.

На рисунках 25 и 26 показаны результаты моделирования в среде Micro-Cap процедуры модуляции. Общий суммарный сигнал поступает в канал связи, после прохождения которого – на фильтр демодулятора. Считаем, что форма сигнала не изменилась. Демодулятор будет выделять свои соответствующие составляющие квадратурных сигналов из суммарного сигнала. Мы получим два сигнала на выходах канального цифрового комплексного полосового КИХ-фильтра демодулятора.

ТАБЛИЦА 1. Пример определения взвешивающих коэффициентов модуляции

TABLE 1. Example of Determining the Weighting Coefficients of Modulation

Тактовый интервал	Бинарная комбинация	$K_{\text{В}}(m)$	$K_{\text{М}}(m)$
0	–	5	0
1	0001	1	3
2	0010	1	–1
3	1100	–3	1

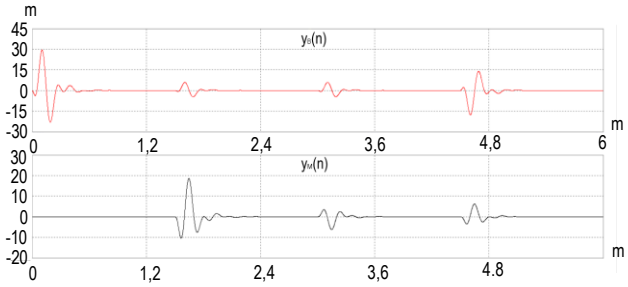


Рис. 25. Выходные сигналы канальных комплексных КИХ-фильтров Баттерворта

Fig. 25. Output Signals of Butterworth Channel Complex FIR-Filters

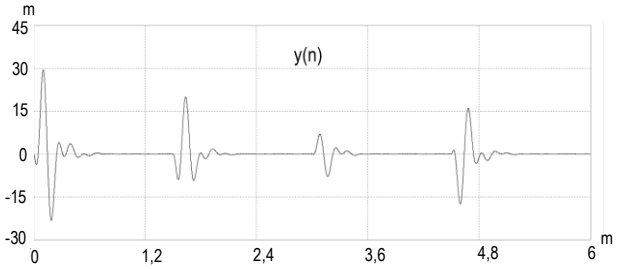


Рис. 26. Суммарный сигнал на выходе модулятора системы передачи данных

Fig. 26. The Total Signal at the Output of the Data Transmission System Modulator

С выхода модулятора сигнал поступает в канал связи, АЧХ которого в рабочем диапазоне частот является неизвестной постоянной величиной K_K, а ФЧХ описывается близкой к линейной функции частоты. Соответственно, характеристика ГВЗ описывается неизвестной постоянной величиной $T_З$. При моделировании канал связи, для которого $K_к = 1$, а $T_з = 0$, будем считать идеальным. На рисунке 27 показаны выходные сигналы комплексного ПФ демодулятора.

$$\begin{array}{llll} 0000, (K_{\text{В}} = 1, K_{\text{М}} = 1); & 0100, (K_{\text{В}} = 3, K_{\text{М}} = 1); & 1000, (K_{\text{В}} = -1), K_{\text{М}} = 1); & 1100, (K_{\text{В}} = -3), K_{\text{М}} = 1); \\ 0001, (K_{\text{В}} = 1, K_{\text{М}} = 3); & 0101, (K_{\text{В}} = 3, K_{\text{М}} = 3); & 1001, (K_{\text{В}} = -1), K_{\text{М}} = 3); & 1101, (K_{\text{В}} = -3), K_{\text{М}} = 3); \\ 0010, (K_{\text{В}} = 1, K_{\text{М}} = -1); & 0110, (K_{\text{В}} = 3, K_{\text{М}} = -1); & 1010, (K_{\text{В}} = -1), K_{\text{М}} = -1); & 1100, (K_{\text{В}} = -3), K_{\text{М}} = -1); \\ 0011, (K_{\text{В}} = 1, K_{\text{М}} = -3); & 0111, (K_{\text{В}} = 3, K_{\text{М}} = -3); & 1011, (K_{\text{В}} = -1), K_{\text{М}} = -3); & 1111, (K_{\text{В}} = -3), K_{\text{М}} = -3). \end{array} \quad (5)$$

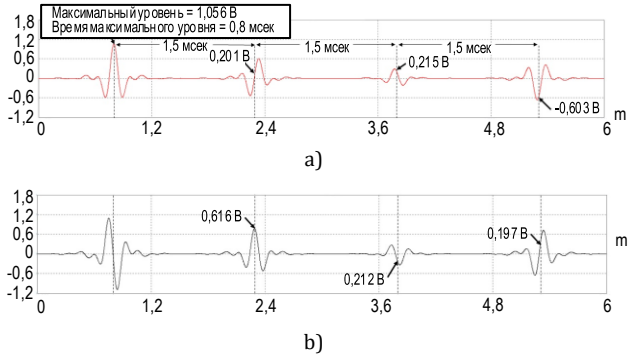


Рис. 27. Выходные сигналы канальных комплексных фильтров Баттерворта демодулятора: a) вещественный выход фильтра; b) мнимый

Fig. 27. Output Signals of the Butterworth channel Complex Filters of the Demodulator: a) Real Filter Output; b) Imaginary

На этапе демодуляции сначала мы определяем максимальный уровень вещественной части сигнала y'_{B1} (1.056 В) и соответствующий ему момент времени (0.8 мс) на опорном тактовом интервале m_1. Затем, чтобы определить значения уровней информационных сигналов на следующих тактовых интервалах, мы добавляем к найденному значению временного отсчетам (0.8 мс) по 1.5 мс на один тактовый интервал. В рассчитанные моменты времени мы определяем значения уровней информационных сигналов (y'_{Bm}, y'_{Mm}) на тактовых интервалах $m = 2, 3, 4$.

Измеряемые значения коэффициентов модуляции $(K'_B(2), K'_M(2))$, $(K'_B(3), K'_M(3))$ и $(K'_B(4), K'_M(4))$ определяем по выражению (6).

$$K'_B(2) = \left(\frac{y'_{B2}}{y'_{B1}}\right) \times K_B(1), K'_M(2) = \left(\frac{y'_{M2}}{y'_{B1}}\right) \times K_B(1), K'_B(3) = \left(\frac{y'_{B3}}{y'_{B1}}\right) \times K_B(1), K'_M(3) =$$
$$= \left(\frac{y'_{M3}}{y'_{B1}}\right) \times K_B(1), K'_B(4) = \left(\frac{y'_{B4}}{y'_{B1}}\right) \times K_B(1) \text{ и } K'_M(4) = \left(\frac{y'_{M4}}{y'_{B1}}\right) \times K_B(1). \quad (6)$$

В нашем случае $K_B(1) = 5$. В таблице 2 приведены границы принятия решений для коэффициентов модуляции, а в таблице 3 – уровни принятых сигналов и значения коэффициентов модуляции для рассматриваемого примера.

ТАБЛИЦА 2. Границы принятия решений

TABLE 2. Decision-Making Boundaries

Нижний порог	Верхний порог	Принятое решение
> –4	< –2	–3
> –2	< 0	–1
> 0	< 2	1
> 2	< 4	3
> 4	< 6	5

ТАБЛИЦА 3. Принятые решения о значениях коэффициентов

TABLE 3. Decisions Made on the Values of Coefficients

Тактовый интервал	y'_{Bm} (В)	y'_{Mm} (В)	$K'_B(m)$	$K'_M(m)$	Принятое решение	
					$K_B(m)$	$K_M(m)$
1	1.056	0	–	–	5	0
2	0.201	0.616	0.95	2.92	1	3
3	0.215	–0.212	0.986	–1.004	1	–1
4	–0.603	0.197	–2.855	0.93	–3	1

По таблице 3 получается принятое решение. Принятая кодовая комбинация совпала с переданной (0001 0010 1100).

Заключение

Разработана схема формирования КАМ-сигнала и схема демодулятора КАМ-сигнала на базе аналоговых комплексных фильтров Баттерворта и обратного цифрового КИХ-фильтра. Методика расчета комплексного ПФ с заданными значениями полюсов НЧ-прототипа Баттерворта позволяет разработать структурные схемы в виде последовательного соединения простых комплексных звеньев с НЧ-прототипом 1-го порядка. Проведено схемотехническое моделирование процедур квадратурной модуляции и демодуляции в среде Micro-Cap. Результаты схемотехнического моделирования подтвердили работоспособность предложенной фильтровой системы передачи данных.

В дальнейшем предполагаются исследования данной системы, но построенной в виде параллельного или параллельно-последовательного соединения простых комплексных звеньев с НЧ-прототипом 1-го порядка.

Список источников

1. Meng M., Wu K.-L. Direct synthesis of general Chebyshev bandpass filters with a frequency variant complex load // Proceedings of the MTT-S International Microwave Symposium (Anaheim, USA, 23–28 May 2010). IEEE, 2010. PP. 433–436. DOI:10.1109/MWSYM.2010.5517836

2. Guest E., Mijatovic N. Discrete-Time Complex Bandpass Filters for Three-Phase Converter Systems // IEEE Transactions on Industrial Electronics. 2019. Vol. 66. Iss. 6. PP. 4650–4660. DOI:10.1109/TIE.2018.2860554
3. Zeidan J., Bila S., Nasser A., Perigaud A., Hamieh A. Quasi-reflection-less bandpass filter with a variable frequency plan // Proceedings of the MTT-S International Microwave Filter Workshop (IMFW, Perugia, Italy, 17–19 November 2021). IEEE, 2021. PP. 8–10. DOI:10.1109/IMFW49589.2021.9642366
4. Балашов В.А., Воробиенко П.П., Ляховецкий Л.М. Системы передачи ортогональными гармоническими сигналами. М.: Эко-Трендз, 2012. 223 с.
5. Гребенко Ю.А., Аунг К.М. Проектирование комплексных полосовых фильтров на базе программируемых аналоговых интегральных схем // Электросвязь. 2020. № 8. С. 71–74. DOI:10.34832/ELSV.2020.9.8.009
6. Гребенко Ю.А., Поляк Р.И. Линеаризация фазочастотной характеристики фильтра нижних частот // Вестник МЭИ. 2015. № 3. С. 90–94.
7. Гребенко Ю. А., Поляк Р. И. Линеаризация фазочастотной характеристики комплексного аналогового полосового фильтра // Вестник МЭИ. 2015. № 4. С. 79–85.
8. Гребенко Ю.А., Чжо Зей Я. Комплексные активные RC-фильтры на идентичных звеньев // Радиотехника. 2008. № 2. С. 61–64
9. Гребенко Ю.А. Методы цифровой обработки сигналов в радиоприемных устройствах. М.: Издательский дом МЭИ, 2006. 48 с.
10. Лайонс Р. Цифровая обработка сигналов. Пер. англ. М.: Бином, 2006. 656 с.

References

1. Meng M., Wu K.-L. Direct synthesis of general Chebyshev bandpass filters with a frequency variant complex load. *Proceedings of the* MTT-S *International Microwave Symposium,* 23–28 May 2010, Anaheim, USA). IEEE; 2010. p.433–436. DOI:10.1109/MWSYM.2010.5517836
2. Guest E., Mijatovic N. Discrete-Time Complex Bandpass Filters for Three-Phase Converter Systems. *IEEE Transactions on Industrial Electronics.* 2019;66(6):4650–4660. DOI:10.1109/TIE.2018.2860554
3. Zeidan J., Bila S., Nasser A., Perigaud A., Hamieh A. Quasi-reflection-less bandpass filter with a variable frequency plan. *Proceedings of the MTT-S International Microwave Filter Workshop, IMFW,* 17–19 November 2021, Perugia, Italy). IEEE, 2021. PP. 8–10. DOI:10.1109/IMFW49589.2021.9642366
4. Balashov V.A., Vorobienko P.P., Lyakhovetsky L.M. *Transmission Systems by Orthogonal Harmonic Signals*. Moscow: Eco-Trends Publ.; 2012. 223 p.
5. Grebenko Yu.A., Aung Ko M. Design of Complex Band-Pass Filters Based on Programmable Analog Integrated Circuits. Electrosvyaz. 2020;8:71–74. DOI:10.34832/ELSV.2020.9.8.009
6. Grebenko Yu.A., Polyak R.I. Linearization Phase Response Lowpass Filter. *Vestnik MEI.* 2015;3:90–94.
7. Grebenko Yu.A., Polyak R.I. Linearization of the Phase-Frequency Characteristic of a Complex Analog Bandpass Filter. *Vestnik MEI.* 2015;4:79–85.
8. Grebenko Yu.A., Chgo Zei Ya. The Complex Active RC-Filters on Identical Units. *Radiotekhnika*. 2008;2:61–64.
9. Grebenko Yu.A. *Methods of Digital Signal Processing in Radio Receivers*. Moscow: MEI Pub.; 2006. 48 p.
10. Lyons R. *Understanding Digital Signal Processing*. New Jersey: Prentice HALL Professional Technical Reference, 2004.

Статья поступила в редакцию 06.07.2023; одобрена после рецензирования 17.07.2023; принята к публикации 25.08.2023.
The article was submitted 06.07.2023; approved after reviewing 17.07.2023; accepted for publication 25.08.2023.

Информация об авторах:

ГРЕБЕНКО Юрий Александрович — доктор технических наук, профессор, профессор кафедры формирования и обработки радиосигналов Национального исследовательского университета «МЭИ»
https://orcid.org/0009-0005-6702-1134

ПОЛЯК Роман Игоревич — кандидат технических наук, доцент кафедры формирования и обработки радиосигналов Национального исследовательского университета «МЭИ»
https://orcid.org/0009-0002-2030-5605

ЯН Лин Пайнг — аспирант кафедры формирования и обработки радиосигналов Национального исследовательского университета «МЭИ»
https://orcid.org/0009-0001-3581-8473

Научная статья
УДК 621.396
DOI:10.31854/1813-324X-2023-9-5-16-24

Анализ энергоэффективности схемы прерывистого приема в системах связи 5G NR

Григорий Александрович Ермолаев ✉, gregory.a.ermolaev@gmail.com
Олеся Викторовна Болховская, obol@rf.unn.ru
Александр Александрович Мальцев, maltsev@rf.unn.ru

Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, 603950, Российская Федерация

Аннотация: *Целью данной работы является анализ схем энергосбережения пользовательского оборудования в первых релизах (Release 15) систем сотовой мобильной связи пятого поколения 5G NR, а также определение возможных направлений повышения энергоэффективности этих систем. В работе описаны выявленные недостатки существующей схемы прерывистого приема, используемой для энергосбережения пользовательского оборудования в 5G NR. Путем имитационного моделирования системы проведен детальный анализ эффективности снижения потребления энергии пользовательским оборудованием при использовании схемы прерывистого приема для различных моделей трафика и ключевого сценария развертывания систем беспроводной сотовой мобильной связи 5G. Анализ результатов моделирования показал, что использование схемы прерывистого приема не позволяет достичь верхней границы возможного энергосбережения для всех исследованных моделей трафика по ряду причин, описанных в данной работе.*

Ключевые слова: *системы радиосвязи, энергосбережение, схема прерывистого приема, DRX, приемо-передающее оборудование, алгоритмы цифровой обработки сигналов, 5G NR*

Финансирование: *исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-32-90197.*

Ссылка для цитирования: Ермолаев Г.А., Болховская О.В., Мальцев А.А. Анализ энергоэффективности схемы прерывистого приема в системах связи 5G NR// Труды учебных заведений связи. 2023. Т. 9. № 5. С. 16–24. DOI:10.31854/1813-324X-2023-9-5-16-24

Energy Efficiency Analysis of the Discontinuous Reception Scheme in 5G NR Communication Systems

Gregory Ermolaev ✉, gregory.a.ermolaev@gmail.com
Olesya Bolkhovskaya, obol@rf.unn.ru
Alexander Maltsev, maltsev@rf.unn.ru

Lobachevsky State University of Nizhny Novgorod,
Nizhny Novgorod, 603950, Russian Federation

Abstract: *The purpose of this work is to analyze the energy saving schemes of user equipment in the first releases (Release 15) of the fifth generation (5G NR) cellular mobile communication systems, as well as to identify possible ways to improve the energy efficiency of these systems. The paper describes the identified drawbacks of the existing discontinuous reception scheme (DRX-scheme) used for user equipment energy saving in 5G NR. The effectiveness of discontinuous reception scheme in reducing user equipment energy consumption was analyzed in detail by simulation study for various traffic models and a key scenario for the deployment of 5G wireless cellular mobile communication*

systems. Analysis of the simulation results shows that the discontinuous reception scheme does not reach the upper limit of possible energy saving for all the traffic models studied for a number of reasons described in this paper.

Keywords: *radio communication systems, energy saving, discontinuous reception scheme, DRX, receiving and transmitting equipment, digital signal processing algorithms, 5G NR*

Funding: *This research was funded by RFBR according to the research project No. 20-32-90197*.

For citation: Ermolaev G., Bolkhovskaya O., Maltsev A. Energy Efficiency Analysis of the Discontinuous Reception Scheme in 5G NR Communication Systems. *Proceedings of Telecommun. Univ.* 2023;9(5):16–24. DOI:10.31854/1813-324X-2023-9-16-24

Введение

Беспроводная мобильная связь является одной из наиболее быстро развивающихся областей современных радиоэлектронных и телекоммуникационных систем. Однако проведенный анализ показывает, что стандарт LTE (*аббр. от англ.* Long Term Evolution) систем мобильной связи четвертого поколения (4G) не способен выполнить требования, предъявляемые потребителями к пропускной способности, времени задержки передачи данных и стабильности подключения к сети [1]. По сравнению со стандартом 4G LTE, стандарт 5G NR призван обеспечить в 20 раз более высокую пиковую скорость передачи данных и в 10 раз меньшую временную задержку [2]. В результате в системах связи пятого поколения при более высокой плотности пользователей будет реализована надежная передача данных за счет использования высокочастотных диапазонов и передовых сетевых технологий.

Для обеспечения высокоскоростной передачи в сетях 5G, несмотря на увеличение спектральной эффективности по сравнению с сетями 4G [3, 4], требуется расширение используемых частотных диапазонов, что, однако, приводит к увеличению потребления энергии как для абонентского терминала пользователя (АТ), так и для базовых станций (БС). Таким образом, помимо новых алгоритмов физического уровня для поддержания высокоскоростной стабильной работы системы связи [5–7], необходимо провести анализ и наметить пути для дальнейшего улучшения используемых в 5G NR Release 15 схем энергосбережения, особенно критичных для АТ [8–10].

Сектор радиосвязи Международного Союза Электросвязи (МСЭ-Р) определяет энергоэффективность как одно из технических требований к производительности систем стандарта мобильной сотовой связи IMT-2020 (Международная мобильная связь 2020). Согласно отчету МСЭ-Р [2], энергоэффективность устройства может быть оценена, исходя из двух критериев:

– эффективности (качества) передачи данных в случае высокой загруженности системы связи;

– низкого энергопотребления при отсутствии передаваемых данных.

Качество передачи данных в случае высокой загруженности системы связи определяется средней спектральной эффективностью. Низкое энергопотребление при отсутствии данных можно оценить по доле времени, в котором пользователь находится в спящем режиме.

Заметная часть потребляемой АТ энергии в 5G NR, как и в 4G LTE, приходится на режим доступа к сети (режим RRC_CONNECTED [1], *аббр. от англ.* Radio Resource Control connected), что обуславливается обработкой всей агрегированной полосы частот, активными электрическими радиочастотными цепями приема/передачи, мониторингом контрольной информации, а также динамическими переходами в энергосберегающий режим и выходами из него. Для обеспечения контроля над энергопотреблением АТ в первом релизе (Release 15) стандарта 5G NR используется внедренная еще в 4G LTE схема прерывистого приема (схема DRX, *аббр. от англ.* Discontinuous Reception) [1]. Выявленные в настоящей работе недостатки существующей схемы энергосбережения позволяют определить основные направления дальнейших исследований для повышения энергоэффективности АТ в последующих релизах стандарта 5G NR.

1. Методы энергосбережения пользовательского оборудования в системах мобильной связи 4G LTE

Трафик пакетов данных часто бывает очень интенсивным, с периодическими интервалами передачи, за которыми следуют более длительные периоды «молчания». С точки зрения задержки передачи данных, полезно отслеживать управляющую сигнализацию нисходящей линии связи (Downlink) от БС к пользователю в каждом временном слоте, чтобы получать гранты (разрешения) для передачи по восходящей линии связи (Uplink) или передачи данных в Downlink и мгновенно реагировать на изменения в поведении трафика. В то же время, такой график мониторинга сопряжен с большими затратами с точки зрения энергопотребления устройства, так как энергозатраты электрической схемы приемника в мобильном устройстве составляют большую часть от его общего энергопотребления.

Алгоритм прерывистого приема

Для снижения энергопотребления мобильного устройства, стандарт 5G NR включает в себя механизм прерывистого приема, следующий той же схеме применения, что и в 4G LTE [1], но с доработками, учитывающими новые нумерологии, поддерживаемые стандартом связи 5G NR.

Основным элементом DRX-схемы является настраиваемый DRX-цикл. В соответствии с DRX-схемой, АТ переходит в активное состояние, включая при этом таймер On duration, с определенной периодичностью, равной длительности DRX-цикла. При сконфигурированном DRX-цикле АТ мобильного устройства отслеживает контрольную информацию, передаваемую в Downlink, только в активном состоянии, переходя в спящий режим, т. е., выключая электрическую схему приемника в оставшееся время (см. рисунок 1), если до истечения длительности работы таймера On duration АТ не получило контрольную информацию от БС.

Это позволяет значительно снизить энергопотребление: чем дольше цикл, тем больший процент времени АТ находится в спящем режиме, т. е. тем ниже среднее энергопотребление в единицу времени. Однако это подразумевает введение ограничений для планировщика передачи данных на БС, поскольку передача для мобильного устройства может быть запланирована и доведена до него только тогда, когда оно активно (включен его приемник) в соответствии с DRX-циклом.

Во многих ситуациях, если передача на мобильное устройство была запланирована и устройство активно принимает или передает данные, весьма вероятно, что в ближайшем будущем будет запланирована следующая передача. Одной из причин может быть то, что было невозможно передать все данные из буфера передачи при использовании одной передачи в Downlink, и, следовательно, необходимы дополнительные передачи.

Ожидание следующего активного периода, т. е. начала следующего цикла, в соответствии с DRX-схемой приведет к дополнительным задержкам. Следовательно, чтобы уменьшить задержки, мобильное устройство пользователя остается в активном состоянии, т. е. с включенным приемником, в течение определенного настраиваемого времени после последней передачи от БС. АТ реализует это, запуская таймер бездействия (*от англ.* Inactivity timer) каждый раз после приема данных, и оставаясь до окончания работы таймера в режиме включенного приемника для мониторинга следующего планирования передачи данных от БС. В связи с тем, что стандартом связи 5G NR поддерживается несколько нумерологий, длительности таймеров DRX-схемы ниже указаны в миллисекундах для отвязки их от определенной нумерологии.

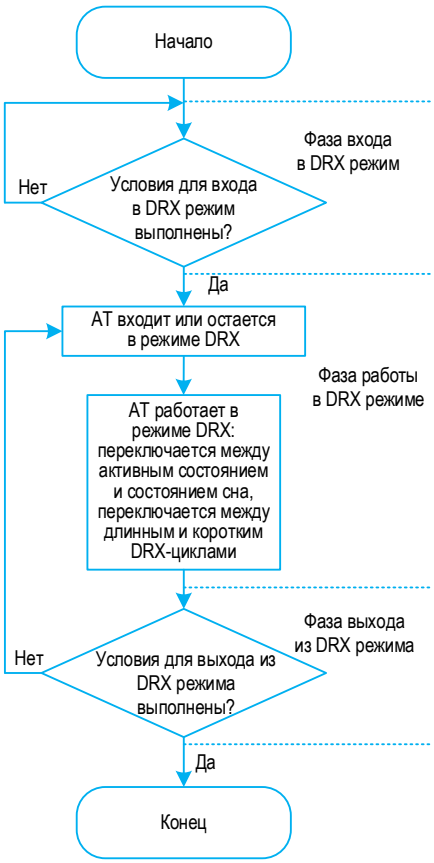


Рис. 1. Блок-схема алгоритма работы схемы прерывистого приема

Fig. 1. Block Diagram of the Operation Algorithm of the Discontinuous Reception Scheme

Проблемные места алгоритма прерывистого приема

На рисунке 2 приведены диаграммы распределения расхода энергии во времени при использовании DRX-схемы (в соответствии со стандартом 5G NR Release 15) и распределения размера буфера данных для передачи на БС. На рисунке 2 показаны основные состояния и работа таймеров DRX-схемы:

1) глубокий сон – неактивное состояние с наименьшим расходом энергии в единицу времени (на рисунке 2 показано желтым цветом); АТ переходит в данное состояние, если временной интервал до следующего перехода в активное состояние больше или равен 20 мс;

2) легкий сон – неактивное состояние с бóльшим расходом энергии в единицу времени, чем в состоянии глубокого сна (на рисунке 2 показано зеленым цветом); АТ переходит в данное состояние, если временной интервал до следующего перехода в активное состояние больше или равен 6 мс и меньше 20 мс;

3) активное состояние – режим, в котором АТ производит мониторинг контрольной информации и/или прием данных; АТ находится в данном режиме во время включенных таймеров On duration (начало каждого DRX-цикла) и Inactivity timer (после каждого приема данных); на рисунке 2

промежутки времени, когда АТ находится в активном состоянии, показаны синим (On duration) и фиолетовым цветом (Inactivity timer).

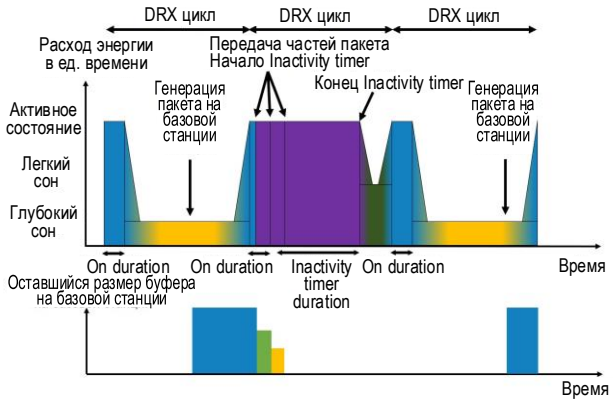


Рис. 2. Распределение расхода энергии во времени при использовании

Fig. 2. Energy Consumption Distribution over Time when Using the Discontinuous Reception Scheme

Как можно увидеть из рисунка 2, за счет периодического перехода в состояние сна, т. е. отключения электрической схемы приемника, значительно сокращается среднее количество потраченной пользователем энергии за единицу времени. Однако из приведенных диаграмм также видно, что в течение каждого DRX-цикла существуют отрезки времени, в течение которых АТ было в активном состоянии, в то время, когда на БС не было ни одного пакета данных для этого пользователя.

Таким образом, из проведенного анализа можно выделить два основных проблемных места DRX-схемы из стандарта 4G LTE:

1) продолжительность работы таймера On duration в случае, когда на БС нет пакета для пользователя перед началом DRX-цикла (первый и третий DRX-циклы);

3) большая продолжительность работы таймера Inactivity после передачи последней части пакета (DRX-цикл 2). В данном случае, если в течение работы таймера Inactivity на БС не появился новый пакет, то за время всей работы таймера пользователь также находится в активном состоянии, в то время как на БС нет пакета для пользователя.

Первый из недостатков DRX-схемы приема, описанных выше, обуславливается тем, что в 5G NR Release 15 АТ пробуждается, т. е. включает электрическую схему приемника в начале каждого DRX-цикла. Таким образом, когда пакеты на БС отсутствуют для определенного АТ, будет тратиться энергия на мониторинг контрольной информации в течение работы таймера On duration. Такие случаи часто встречаются в случае разреженного трафика.

Второй недостаток схемы прерывистого приема из описанных выше оказывает большее негативное воздействие на энергоэффективность АТ по причине относительно большей длительности работы Inactivity timer по сравнению с таймером On duration. Например, для стандартных параметров модели трафика FTP 3 (*аббр. от англ.* File Transfer Protocol 3) оптимальная (с точки зрения сохранения задержки передачи пакета) конфигурация DRX-цикла включает в себя длительности таймеров On duration и Inactivity равные 8 мс и 100 мс, соответственно. Поэтому в течение работы таймера Inactivity АТ будет терять много энергии.

2. Описание системной модели сети связи

Моделирование системы связи 5G для изучения схем энергосбережения АТ проводилось с использованием симулятора системного уровня, основанного на стандарте 5G NR Release 15 [11–14], в котором можно выделить следующие основные блоки:

1) развертывание сети – расположение БС согласно типовой структуре сети мобильной сотовой связи, генерация координат пользователей согласно равномерному распределению;

2) генерация канала связи для процедур ассоциации пользователей к БС, вычисления пользователем информации о состоянии канала (CSI, *аббр. от англ.* Channel State Information) для передачи на БС и для процедуры приема сигнала пользователем в течение всего времени симуляции системы связи;

3) ассоциация мобильных устройств к БС и инициализация связи с БС для каждого АТ;

4) генерация обратной связи о состоянии канала (*от англ.* CSI feedback) от пользователя к БС;

5) блок обработки состояния пользователя согласно алгоритму работы DRX-схемы;

6) планировщик передач на БС, выбирающий пользователя (или группу пользователей) для MU-MIMO (*от англ.* Multi User Multiple-Input and Multiple-Output) схемы передачи данных, оптимальный ранг диаграммы направленности (количество пространственных каналов передачи) и оптимальную конфигурацию диаграммы направленности (весовые коэффициенты для антенной решетки), используя алгоритм пропорционального справедливого (*от англ.* Proportional Fair Algorithm) распределения частотно-временных ресурсов;

7) блок учета интерференции между БС;

8) MMSE (*аббр. от англ.* Minimum Mean Square Error) приемник;

9) абстракция оценки канала связи для имитации ошибки при его оценивании;

10) абстракция канального кодирования/декодирования на физическом уровне (*от англ.* Physical Layer Abstraction) для имитации схемы помехоустойчивого кодирования с использованием кодов с малой плотностью проверки на четность (*от англ.* LDPC codes, Low-Density Parity-Check codes) из стандарта связи NR.

Сценарии моделирования беспроводных мобильных систем связи

При разработке новых релизов стандарта 5G NR используются три ключевых сценария, принятых консорциумом 3GPP (*аббр. от англ.* 3rd Generation Partnership Project) для моделирования системы связи 5G на системном уровне [15]:

1) UMi (*аббр. от англ.* Urban Micro) – сценарий моделирования открытой городской местности размера порядка 200 м;

2) UMa (*аббр. от англ.* Urban Macro) – сценарий моделирования города с БС, смонтированными над уровнями крыш окружающих зданий;

3) Indoor – сценарий моделирования работы системы связи, предназначенный для описания различных типичных сценариев развертывания сети связи внутри помещений, включая офисные помещения и торговые центры.

При анализе DRX-схемы в данной работе моделирование проводилось для UMa-сценария, так как он описывает наиболее часто встречающийся в реальности сценарий развертывания систем беспроводной мобильной связи в больших городах.

Модели трафика мобильных систем связи

Перед началом разработки систем связи 4G LTE модели трафика были представлены группой TSG-RAN WG1 (*аббр. от англ.* Technical Specification Group Radio Access Network Working Group 1) консорциума 3GPP в документе [16], который содержит конкретные примеры и соответствующие параметры для каждой модели. Этот документ включает описание комплексных моделей трафика для наиболее важных сервисов, таких как FTP, Web-browsing (просмотр веб-страниц), Video streaming (потоковое видео), Gaming (игровой трафик) и модель передачи голосового трафика VoIP (*аббр. от англ.* Voice over Internet Protocol).

Рассмотрение каждой модели трафика одинаково важно при анализе энергоэффективности АТ, так как каждый тип трафика составляет заметную часть в общем объеме. В таблице 1 показано примерное процентное распределение пакетов пользователей (по разным типам трафика) в реальных системах беспроводной мобильной связи, полученное на основе анализа реальных данных [16].

ТАБЛИЦА 1. Распределение пакетов пользователей по типам трафика

TABLE 1. Percentage Distribution of User Packets by Traffic Types

Модель	Категория трафика	Процент трафика
FTP	Сервис с негарантированной доставкой	10 %
VoIP	Трафик с передачей в реальном времени	30 %
Web Browsing	Интерактивный	20 %
Video Streaming	Потоковая передача данных	20 %
Gaming	Интерактивный трафик с передачей в реальном времени	20 %

При моделировании системы связи 5G на системном уровне использовались значения параметров моделей трафика из [16], за исключением модели трафика FTP 3, для которой в начале разработки стандарта NR консорциумом 3GPP был выбран фиксированный размер пакетов, равный 0.5 Мбайт, со средним временем между приходами пакетов – 0.2 с.

3. Результаты компьютерного моделирования. Анализ энергоэффективности схемы прерывистого приема

Эффективность использования DRX-схемы для сокращения энергии, потребляемой АТ в системах беспроводной связи стандарта 5G NR Release 15, была проанализирована путем компьютерного моделирования на системном уровне. Моделирование проводилось для передачи в Downlink в частотном диапазоне FR1 (*аббр. от англ.* Frequency Range 1 [17]). Были выбраны стандартные параметры для данного сценария, такие как: несущая частота 4 ГГц, ширина полосы частот 100 МГц, расстояние между поднесущими 30 кГц. Использовалось расстояние между БС, равное 200 м для моделирования плотного расположения БС в UMa-сценарии.

Параметры моделирования на системном уровне представлены в таблице 2 (в соответствии с принятыми моделями канала и обозначениями из [15]).

ТАБЛИЦА 2. Параметры моделирования

TABLE 2. Simulation Parameters

Параметр	Значение
Несущая частота	4 ГГц
Сценарий	Dense Urban Macro FR1 DL
Ширина полосы частот	100 МГц
Расстояние между поднесущими	30 кГц
Модель канала связи	IMT UMa A
Расстояние между БС	200 м
Антенная конфигурация БС (*Mg*, *Ng*, *M*, *N*, *P*; *Mp*, *Np*)	(1, 1, 8, 8, 2; 2, 8) 32 приемного/передающего элемента
Расстояние между антеннами БС (dV, dH) × λ	(0.8, 0.5)
Антенная конфигурация АТ (*Mg*, *Ng*, *M*, *N*, *P*; *Mp*, *Np*)	(1, 1, 1, 2, 2; 1, 2) 4 приемных/передающих элементов
Расстояние между антеннами АТ (dV, dH) × λ	(0.5, 0.5)
Конфигурация аналоговой диаграммы направленности БС	Одинарный луч в направлении 102 ° от вертикального направления

Параметр	Значение
Распределение АТ	80 % внутри помещения; 20 % вне помещения
Наивысшая модуляция	256 QAM
Схема канального кодирования	LPDC из стандарта 5G NR Release 15
Схема приема/передачи	Многопользовательская MIMO-схема: до 12 пространственных потоков на БС; до 4 пространственных потоков на АТ
Планировщик	Выбор пользователя, оптимального ранга (количества пространственных каналов передачи) и оптимальной конфигурации диаграммы направленности, используя многопользовательский алгоритм пропорционального справедливого распределения обслуживания
Схема обратной связи	Type II CSI

Для моделирования потребляемой пользователем энергии были определены типы временных слотов в зависимости от операции [18], которую пользователь производит в определенном слоте, или, другими словами, в зависимости от состояния, в котором пользователь находится в определенном слоте.

Модель потребляемой пользователем энергии для каждого состояния представлена в таблице 3.

ТАБЛИЦА 3. Модель потребляемой энергии

TABLE 3. Energy Consumption Model

Состояние пользовательского оборудования	Потребляемая энергия за временной слот, усл. ед.
Активное состояние (мониторинг контрольной информации)	100
Активное состояние (прием пользователем данных)	300
Режим микросна	45
Режим легкого сна	20
Переход в легкий сон	100 (за всю длительность перехода)
Режим глубокого сна	1
Переход в глубокий сон	450 (за всю длительность перехода)

Значения потребляемой энергии за временной слот в данной модели представлены в условных единицах после нормировки на энергию, потребляемую в режиме глубокого сна. Таким образом, потребляемая энергия за временной слот в течение режима глубокого сна равна единице.

При моделировании для различных моделей трафика, описанных в [16], использовались определенные оптимизированные конфигурации [18] DRX-схемы, представленные в таблице 4, определяющие длительности разных состояний АТ согласно DRX-схеме.

ТАБЛИЦА 4. Конфигурации DRX-схемы

TABLE 4. Configurations of DRX-Scheme

Модель трафика	Длительность, мс		
	DRX цикл	Inactivity timer	On duration
FTP model 3	160	100	8
Instant Messaging	320	80	10
VoIP	40	10	4
Gaming	40	10	4
Web browsing	320	80	10
Video streaming	40	10	4

При моделировании на системном уровне, как и в реальном сценарии развертывания систем связи, пользователи могут находиться в разных условиях, определяемых:

1) загруженностью БС соседних сот; это определяет уровень межсотовой интерференции (*от англ.* Inter-Cell Interference) и, следовательно, отношение сигнал/помеха плюс шум (ОСПШ, *аббр. от англ.* SINR, Signal-to-Interference-plus-Noise Ratio) на приемнике АТ;

2) расстоянием между АТ и обслуживающей БС по причине квадратичной зависимости в свободном пространстве потерь мощности сигнала от расстояния между источником и приемником; также более удаленные от обслуживающей БС пользователи оказываются ближе к БС соседних сот, что также увеличивает уровень межсотовой интерференции;

3) расположением АТ внутри или вне помещения; для АТ, расположенного внутри помещения, значительное ослабление ОСПШ принятого сигнала вызвано потерями при прохождении сигнала внутрь помещения.

По указанным причинам результаты моделирования энергоэффективности DRX-схемы далее представлены отдельно для трех групп пользователей, для которых ОСПШ соответствует 5 %, 50 % и 95 % уровням интегральной функции распределения ОСПШ, представленной на рисунке 3.

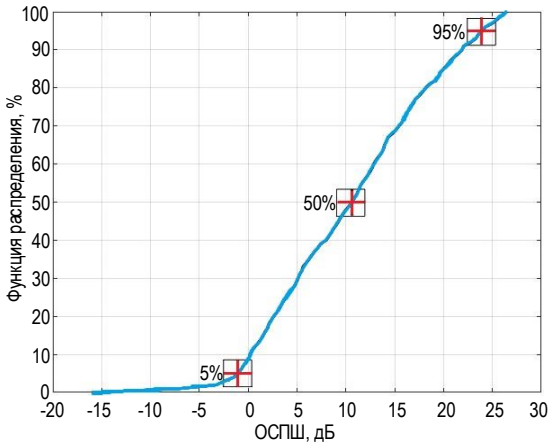


Рис. 3. Функция распределения ОСПШ

Fig. 3. SINR Distribution Function

На рисунке 3 показана полученная путем имитационного моделирования рассматриваемой системы связи интегральная функция распределения так называемого геометрического ОСПШ, при вычислении которого в мощности принятого сигнала не учитывается усиление диаграмм направленности антенных решеток БС и АТ.

В таблице 5 приведены результаты компьютерного моделирования системы связи 5G на системном уровне для всех представленных моделей трафика.

ТАБЛИЦА 5. Выигрыш в энергосбережении при использовании DRX-схемы, %

TABLE 5. Energy Saving Gain when Using the DRX Scheme

ОСПШ	Выигрыш в энергосбережении, %					
	FTP3	IM	VoIP	Gaming	Web browsing	Video streaming
5 %	51	92	71	65	95	57
50 %	55	92	71	65	95	57
95 %	56	92	71	65	95	58

Полученный путем имитирования работы системы связи выигрыш в энергосбережении АТ за счет использования DRX-схемы представлен для трех основных точек интегральной функции распределения ОСПШ (см. рисунок 3). Выигрыш рассчитывается в процентном отношении и равен уменьшению среднего значения потребляемой энергии в единицу времени при использовании DRX-схемы по сравнению с конфигурацией работы АТ без использования DRX-схемы, т. е. по сравнению с АТ, находящимся постоянно в активном состоянии.

Для проведения более детального анализа энергоэффективности системы в дополнение к параметрам каждой модели трафика и конфигурации DRX-цикла целесообразно использовать информацию о загруженности частотно-временных ресурсов на БС при моделировании различных моделей трафика, что позволяет оценить верхнюю границу возможного сокращения потребляемой АТ энергии. Эта оценка может быть получена путем вычисления среднего процента неиспользованных частотно-временных ресурсов системы связи без применения DRX-схемы.

В таблице 6 представлены доли использованных (строка 1) и неиспользованных ресурсов (строка 2) на БС для всех анализируемых в данной работе моделей трафика. В третьей строке таблицы приведена оценка возможного дополнительного сокращения потребляемой АТ энергии, не достигаемое при использовании DRX-схемы. Значения оценки получены для АТ с медианным (50 %) уровнем ОСПШ путем вычитания из среднего процента неиспользованных частотно-временных ресурсов (строка 2 таблицы 6) достигаемого выигрыша в энергосбережении при использовании DRX-схемы (см. строку 2 таблицы 5).

ТАБЛИЦА 6. Доли использованных/неиспользованных частотно-временных ресурсов

TABLE 6. Percentage of Used/Unused Time-Frequency Resources

	FTP3	IM	VoIP	Gaming	Web browsing	Video streaming
Загруженность ресурсов на БС, %	28	1	3.5	11	< 1	28
Средняя доля неиспользованных частотно-временных ресурсов, %	72	99	96.5	89	>99	72
Сокращение потребляемой АТ энергии, недостигаемое схемой DRX, %	~16	~7	~26	~24	~5	~14

Возможное сокращение потребляемой АТ энергии, не достигаемое при использовании DRX-схемы и приведенное в строке 3 таблицы 6, обусловлено неоптимальным распределением интервалов времени в DRX-цикле (см. рисунок 2), и связано, в основном, с интервалами времени, когда АТ находится в активном состоянии, и в то же время на БС отсутствуют пакеты для данного АТ.

Таким образом, для улучшения энергоэффективности АТ в последующих релизах систем связи 5G NR является разработка методов адаптации DRX-схемы за счет использования передачи от БС пользователю определенной контрольной информации, сигнализирующей о состоянии буфера данных на БС, для предотвращения перехода АТ в активное состояние или для перехода АТ в состояние сна, когда на БС отсутствуют пакеты для данного АТ.

Заключение

В статье проведен подробный анализ работы DRX-схемы, используемой для энергосбережения АТ в первом релизе систем сотовой мобильной связи 5G NR Release 15.

В ходе работы были реализованы сценарии моделирования систем беспроводной мобильной связи и модели трафика данных, принятые консорциумом 3GPP. Путем имитационного моделирования системы, работающей по стандарту 5G NR Release 15, проведен детальный анализ эффективности снижения потребления энергии АТ при использовании DRX-схемы для различных моделей трафика и наиболее важного UMa сценария (Dense Urban Macro) развертывания систем беспроводной сотовой мобильной связи 5G в городских условиях.

Анализ результатов моделирования показал, что использование DRX-схемы не позволяет достичь верхней границы возможного энергосбережения для исследованных моделей трафика по причине неоптимального распределения интервалов времени различных состояний пользователя в DRX-цикле, и это связано, в основном, с интервалами времени, когда АТ находится в активном состоянии, в то время как на БС отсутствуют пакеты для данного пользователя.

Направлением повышения энергоэффективности АТ в системах связи 5G NR следующих релизов является введение дополнительной контрольной информации, передаваемой от БС пользователям, для оптимизации распределения временных интервалов различных состояний АТ в DRX-цикле.

Список источников

1. Dahlman E., Parkvall S., Skold J. 5G NR: The Next Generation Wireless Access Technolog. Academic Press, 2018.
2. Rec. ITU-R M.2410-0 (2017). Minimum requirements related to technical performance for IMT-2020 radio interface(s).
3. Mondal B., Sergeev V., Sengupta A., Ermolaev G., Davydov A., Kwon E., et al. MU-MIMO and CSI Feedback Performance of NR/LTE // Proceedings of the 53rd Annual Conference on Information Sciences and Systems (CISS, Baltimore, USA, 20–22 March 2019). IEEE, 2019. PP. 1–6. DOI:10.1109/CISS.2019.8692922
4. Бурков А.А., Тюрликов А.М. Верхняя оценка спектральной эффективности для систем с гибридной решающей обратной связью при ограничении на вид модуляции // Вопросы радиоэлектроники. Серия: Техника телевидения. 2020. № 1. С. 74–83.
5. Ermolaev G.A, Bolkhovskaya O.V., Maltsev A.A. Advanced Approach for TX Impairments Compensation Based on Signal Statistical Analysis at the RX Side // Proceedings of the International Scientific Conference on Wave Electronics and its Application in Information and Telecommunication Systems (WECONF, St. Petersburg, Russia, 31 May 2021 – 04 June 2021). IEEE, 2021. PP. 1–5. DOI:10.1109/WECONF51603.2021.9470687
6. Болховская О.В., Ермолаев Г.А., Трушков С.Н., Мальцев А.А. Прототип приемопередающего оборудования скоростной передачи данных в частотном диапазоне 57–64 ГГц // Труды учебных заведений связи. 2023. Т. 9(2). С. 23–39. DOI:10.31854/1813-324X-2023-9-2-23-39
7. Burkov A., Shneer S., Turlikov A. An Achievability Bound of Energy Per Bit for Stabilized Massive Random Access Gaussian Channel // IEEE Communications Letters. 2020. Т. 25. № 1. PP. 299–302. DOI:10.1109/LCOMM.2020.3023461
8. Herreria-Alonso S., Rodriguez-Perez M., Fernandez-Veiga M., Lopez-Garcia C. Adaptive DRX Scheme to Improve Energy Efficiency in LTE Networks with Bounded Delay // IEEE Journal on Selected Areas in Communications. 2015. Vol. 33. Iss. 12. PP. 2963–2973. DOI:10.1109/JSAC.2015.2478996
9. Wang H.C., Tseng C.C., Chen G.Y., Kuo F.C., Ting K.C. Power saving by LTE DRX mechanism using a mixture of short and long cycles // Proceedings of the International Conference of IEEE Region 10 (TENCON 2013, Xi'an, China, 22–25 October 2013). IEEE, 2013. PP. 1–6. DOI:10.1109/TENCON.2013.6719041
10. Arunagiri P., Nagarajan G. Optimization of power saving and Latency in LTE network using DRX mechanism // Proceedings of the 10th International Conference on Intelligent Systems and Control (ISCO, Coimbatore, India, 07–08 January 2016). IEEE, 2016. PP. 1–4. DOI:10.1109/ISCO.2016.7727036
11. ETSI TS 138.211 (2020-07). NR; Physical channels and modulation (3GPP TS 38.211 version 16.2.0 Release 16).
12. ETSI TS 138.212 (2018-07). NR; Multiplexing and channel coding (3GPP TS 38.212 version 15.2.0 Release 15).
13. ETSI TS 138.213 (2018-07). NR; Physical layer procedures for control (3GPP TS 38.213 version 15.2.0 Release 15).
14. ETSI TS 138.214 (2018-10). NR; Physical layer procedures for data (3GPP TS 38.214 version 15.3.0 Release 15).
15. ETSI TS 138.901 (2020-11). Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 16.1.0 Release 16).
16. 3GPP R1-070674. LTE Physical Layer Framework for Performance Verification. 2007.
17. ETSI TS 138.101-1 (2018-10). 5G; NR; User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone (3GPP TS 38.101-1 version 15.3.0 Release 15).
18. ATIS.3GPP.38.840.V1600 (2019-06). Technical Specification Group Radio Access Network; NR; Study on User Equipment (UE) power saving in NR (Release 16).

References

1. Dahlman E., Parkvall S., Skold J. *5G NR: The Next Generation Wireless Access Technology.* Academic Press; 2018.
2. Rec. ITU-R M.2410-0. *Minimum requirements related to technical performance for IMT-2020 radio interface(s).* 2017.
3. Mondal B., Sergeev V., Sengupta A., Ermolaev G., Davydov A., Kwon E., et al. MU-MIMO and CSI Feedback Performance of NR/LTE. *Proceedings of the 53rd Annual Conference on Information Sciences and Systems, CISS, 20–22 March 2019, Baltimore, USA*. IEEE; 2019. p.1–6. DOI:10.1109/CISS.2019.8692922
4. Burkov A.A., Turlikov A.M. An Upper Estimate of the Spectral Efficiency for Systems with Hybrid Automatic Repeat Request with a Restriction on the Type of Modulation. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2020;1:74–83.
5. Ermolaev G.A, Bolkhovskaya O.V., Maltsev A.A. Advanced Approach for TX Impairments Compensation Based on Signal Statistical Analysis at the RX Side. Proceedings of the International Scientific Conference on Wave Electronics and its Application in Information and Telecommunication Systems, WECONF, 31 May 2021 – 04 June 2021, St. Petersburg, Russia. IEEE; 2021. p.1–5. DOI:10.1109/WECONF51603.2021.9470687
6. Bolkhovskaya O., Ermolaev G., Trushkov S., Maltsev A. Prototype of High-Speed Data Transmission Receiving and Transmitting Equipment in the 57–64 GHz Frequency Range. *Proceedings of Telecommun. Univ*. 2023;9(2):23–39. DOI:10.31854/1813-324X-2023-9-2-23-39

7. Burkov A., Shneer S., Turlikov A. An Achievability Bound of Energy Per Bit for Stabilized Massive Random Access Gaussian Channel. *IEEE Communications Letters*. 2020;25(1):299–302. DOI:10.1109/LCOMM.2020.
8. Herreria-Alonso S., Rodriguez-Perez M., Fernandez-Veiga M., Lopez-Garcia C. Adaptive DRX Scheme to Improve Energy Efficiency in LTE Networks with Bounded Delay. *IEEE Journal on Selected Areas in Communications*. 2015;33(12):2963–2973. DOI:10.1109/JSAC.2015.2478996
9. Wang H.C., Tseng C.C., Chen G.Y., Kuo F.C., Ting K.C. Power saving by LTE DRX mechanism using a mixture of short and long cycles. *Proceedings of the International Conference of IEEE Region 10, TENCON 2013, 22–25 October 2013, Xi'an, China*. IEEE; 2013. p.1–6. DOI:10.1109/TENCON.2013.6719041
10. Arunagiri P., Nagarajan G. Optimization of power saving and Latency in LTE network using DRX mechanism. *Proceedings of the 10th International Conference on Intelligent Systems and Control, ISCO, 07–08 January 2016, Coimbatore, India*. IEEE; 2016. p.1–4. DOI:10.1109/ISCO.2016.7727036
11. ETSI TS 138.211. *NR; Physical channels and modulation (3GPP TS 38.211 version 16.2.0 Release 16)*. July 2020.
12. ETSI TS 138.212. *NR; Multiplexing and channel coding (3GPP TS 38.212 version 15.2.0 Release 15)*. July 2018.
13. ETSI TS 138.213. *NR; Physical layer procedures for control (3GPP TS 38.213 version 15.2.0 Release 15).* July 2018.
14. ETSI TS 138.214. *NR; Physical layer procedures for data (3GPP TS 38.214 version 15.3.0 Release 15)*. October 2018.
15. ETSI TS 138.901. *Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 16.1.0 Release 16)*. November 2020.
16. 3GPP R1-070674. *LTE physical layer framework for performance verification*. 2007.
17. ETSI TS 138.101-1. *5G; NR; User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone (3GPP TS 38.101-1 version 15.3.0 Release 15).* October 2018.
18. ATIS.3GPP.38.840.V1600. *Technical Specification Group Radio Access Network; NR; Study on User Equipment (UE) power saving in NR (Release 16)*. June 2019.

Статья поступила в редакцию 01.08.2023; одобрена после рецензирования 31.10.2023; принята к публикации 07.11.2023.
The article was submitted 01.08.2023; approved after reviewing 31.10.2023; accepted for publication 07.11.2023.

Информация об авторах:

ЕРМОЛАЕВ Григорий Александрович
аспирант кафедры статистической радиофизики и мобильных систем связи Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского
https://orcid.org/0000-0003-4213-953X

БОЛХОВСКАЯ Олеся Викторовна
кандидат физико-математических наук, доцент, доцент кафедры статистической радиофизики и мобильных систем связи Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского
https://orcid.org/0000-0002-6679-9295

МАЛЬЦЕВ Александр Александрович
доктор физико-математических наук, профессор, заведующий кафедрой статистической радиофизики и мобильных систем связи Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского
https://orcid.org/0000-0001-8694-0033

Научная статья
УДК 654.165
DOI:10.31854/1813-324X-2023-9-5-25-34

Алгоритм автоматического размещения базовых станций транкинговых систем связи

Вячеслав Сергеевич Иванов, ivanovmirea1@yandex.ru
Увайсов Сайгид Увайсович, uvajsov@mirea.ru
Иванов Илья Александрович, ivanov_ia@mirea.ru

МИРЭА – Российский технологический университет,
Москва, 119454, Российская Федерация

Аннотация: *Для повышения эффективности процесса проектирования транкинговых систем связи предлагается оригинальный алгоритм автоматического размещения базовых станций с учетом реальных условий распространения радиоволн. Проведенный анализ показал, что существующие методы позволяют размещать их вручную, с заданным шагом или автоматически, но без учета распространения радиоволн, что приводит к некачественному покрытию требуемой территории связью. Целью работы является снижение трудоемкости проектирования транкинговой системы связи, гарантированно покрывающей заданную зону обслуживания. Предлагается рассчитывать зоны обслуживания базовых станций с помощью статистической модели, а также с учетом дополнительных потерь на дифракцию на трассе распространения радиосигнала. Решение задачи автоматического размещения основано на использовании предложенной модифицированной модели Хата, учитывающей рельеф местности в местах нахождения базовой и портативной станций, а также на международной рекомендации по прогнозированию распространения сигнала на конкретной трассе для наземных служб. При автоматическом размещении выполняется критерий оптимальности, заключающийся в размещении минимального числа базовых станций, необходимых для покрытия требуемой территории. Элементом новизны представленного решения является то, что на первом этапе расчета определяется зона обслуживания базовой станции статистическим методом, на втором происходит уточненный расчет зоны обслуживания с учетом реальных условий распространения радиоволн на месте привязки, а на третьем – определяется зона гарантированного обслуживания абонентов. Использование предлагаемого алгоритма позволяет снизить трудоемкость процесса проектирования системы и определить достаточное число базовых станций, необходимых для покрытия заданной территории. Для применения алгоритма в реальной практике разработано соответствующее программное обеспечение. В подтверждение эффективности в работе приведены сравнительные расчеты затрат времени на проектирование транкинговой системы связи с применением разработанного программного обеспечения и широко применяемой программы RadioMobile.*

Ключевые слова: *транкинговые системы связи, зоны обслуживания, базовые станции, проектирование сетей, автоматическое размещение*

Ссылка для цитирования: Иванов В.С., Увайсов С.У., Иванов И.А. Алгоритм автоматического размещения базовых станций транкинговых систем связи // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 25–34. DOI:10.31854/1813-324X-2023-9-5-25-34

Automatic Placement Algorithm of Base Stations Trunking Communication Systems

Vyacheslav Ivanov, ivanovmirea1@yandex.ru
Saigid Uvajsov, uvajsov@mirea.ru
Ilya Ivanov, ivanov_ia@mirea.ru

MIREA – Russian Technological University,
Moscow, 119454, Russian Federation

Abstract: *to increase the efficiency of the trunking communication systems design process, an original algorithm for the automatic placement of base stations is proposed, taking into account the real conditions of radio wave propagation. The analysis showed that the existing methods allow you to place base stations manually, with a given step or automatically, but without taking into account the propagation of radio waves, which leads to poor-quality coverage of the required area with communication. The aim of the work is to reduce the complexity of designing a trunking communication system that is guaranteed to cover a given service area. It is proposed to calculate the service areas of base stations using a statistical model, as well as taking into account additional diffraction losses on the radio signal propagation path. The solution to the problem of automatic placement of base stations is based on the use of the proposed modified Hut model, which takes into account the terrain at the locations of the base and portable stations, as well as on the international recommendation for predicting signal propagation on a specific route for ground services. With automatic placement, the optimality criterion is fulfilled, which consists in placing the minimum number of base stations necessary to cover the required territory. An element of the novelty of the presented solution is that at the first stage of the calculation, the service area of the base station is determined by the statistical method, at the second stage, an updated calculation of the service area takes place taking into account the actual conditions of radio wave propagation at the binding site, and at the third stage, the zone of guaranteed customer service is determined. The use of the proposed algorithm makes it possible to reduce the complexity of the system design process and determine the sufficient number of base stations needed to cover a given area with communication. Appropriate software has been developed to apply the algorithm in real practice. To confirm the effectiveness of the algorithm, the paper presents comparative calculations of the time spent on designing a trunking communication system using the developed software and using the widely used RadioMobile program.*

Keywords: *trunking communication systems, service areas, base stations, network design, automatic placement*

For citation: Ivanov V., Uvajsov S., Ivanov I. Automatic Placement Algorithm of Base Stations Trunking Communication Systems. *Proceedings of Telecommun. Unive.* 2023;9(5):25–34. DOI:10.31854/1813-324X-2023-9-5-25-34

Введение

Транкинговые системы (ТС) – это один из видов систем подвижной связи наравне с сотовыми системами, системами персонального радиовызова и спутниковой связью. Особенность ТС заключается в автоматическом распределении ограниченного количества каналов между абонентами. Если сравнивать ТС с конвенциональными системами, то в последних существует проблема неравномерного использования каналов связи. За каждым каналом связи закрепляются определенные абоненты, которые находятся в ожидании, если канал занят, при свободных соседних каналах.

Отсюда появилась вышеуказанная проблема, которая в ТС не наблюдается. На данный момент территория нашей страны не покрыта связью полностью, т. к. операторы мобильной связи не устанавливают базовые станции (БС) в малонаселенных местах из-за низкой прибыли. Поэтому ТС используют крупные компании, занимающиеся добычей полезных ископаемых там, где нет мобильной связи. За счет высокой скорости установления соединения ТС используют службы спасения, а также службы безопасности, как государственные, так и коммерческие [1].

Расчет зон обслуживания БС производится с использованием статистических и детерминированных методов. В статистических методах моделируется окружность, на границах которой считается, что обеспечивается качественная связь. Известно большое количество статистических моделей, среди которых модель Окумуры, модель Хата и др., с помощью которых можно рассчитать зону обслуживания БС. В детерминированных учитываются принципы теории распространения радиоволн и зона обслуживания имеет сложную геометрическую форму [2]. На рисунке 1 окружностью показана зона обслуживания БС, рассчитанная статистическим методом. На этом же рисунке в форме «амебы» представлена зона обслуживания БС, рассчитанная с учетом реальных условий распространения радиоволн.

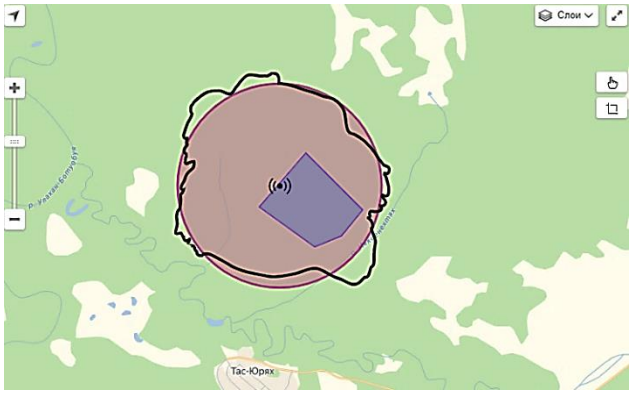


Рис. 1. Зоны обслуживания БС

Fig. 1. BS Service Areas

Проектирование ТС связи ведется в различных программах, как отечественных (RadioPlanner), так и зарубежных (RadioMobile), которые позволяют с высокой точностью рассчитать зоны обслуживания БС. В данных программах пользователю необходимо указать технические характеристики приемопередающего оборудования, антенн и места установки БС. Если перед инженером-проектировщиком стоит задача минимизировать количество БС с целью сокращения финансовых затрат на развертывание ТС, то ему необходимо потратить определенное количество времени на подбор оптимальных мест установки БС.

В работе Л.Л. Егорова и соавторов [3] предложен алгоритм, который на основе метода наименьших квадратов осуществляет автоматический просчет кластерной системы БС сотовой связи. Данный алгоритм позволяет получить для проектировщика различные варианты решения для кластера различной размерности. При этом предполагается, что места расположений БС известны, и задача сводится к геометрическому варианту алгоритма расчета зон покрытия БС.

Работа Р.Р. Мухаджинова [4] посвящена применению генетического алгоритма в размещении БС, где на первом этапе происходит поиск места расположения, при котором в зоне обслуживания площадь теневых зон будет минимальная. Далее производится поиск теневых зон, обусловленных интерференцией основного сигнала с отраженными сигналами от подстилающей поверхности и объектов застройки. Затем проводится поиск оптимального расположения дополнительных ретрансляторов, обеспечивающих приемлемый уровень сигнала приемопередающих станций в теневых зонах.

В работе [5] К.А. Павловской решается задача размещения минимального количества БС на заданной территории с помощью генетического алгоритма. Сначала определяются различные варианты расстановки БС, далее производится выборка случайного числа БС1 и БС2 – наиболее «сильных хромосом» из общего количества. Затем начинается главный цикл, где происходит оценка родительских хромосом, с учетом расстояния радиуса $R1$ и $R2$ от каждой БС до каждого абонента.

В работе А.А. Мухтарова и О.Ю. Першина [6] предложены алгоритмы построения графов информационных потоков, позволившие формализовать задачи в виде соответствующих моделей математического программирования. Задача оптимального размещения БС решается с целью минимизации их общей стоимости, путем перебора возможных мест расположения.

В работе Р.Р. Аминовой [7] описан разработанный ею алгоритм определения расположения БС, который проходится по точкам координатной сетки, наложенной на обслуживаемую территорию. Алгоритм начинает перебор точек и по итогам опроса всех точек размещает окружность в «наилучшей» точке. При этом все точки, которые находятся внутри окружности, не учитываются при дальнейших расчетах. Далее алгоритм повторяет свою работу до тех пор, пока не будут покрыты все точки.

Работа Е.С. Скакова и В.Н. Малыш [8] посвящена разработке модифицированных алгоритмов вероятностного поиска с запретами и мультистарта. Алгоритм мультистарта позволяет получить локальные решения, которые являются наилучшими на данный момент, что впоследствии приведет к оптимальному решению глобальной задачи.

Анализ данных работ показал, что автоматическое размещение БС с учетом реальных условий распространения радиоволн является актуальной задачей, позволяющей снизить трудоемкость проектирования ТС, гарантированно покрывающих заданную зону обслуживания.

Постановка задачи

Для расчета зон обслуживания БС по статистическим методам необходимо определить технические характеристики приемопередающего оборудования, антенн, а также статистическую модель для расчета. Алгоритм, представленный в данной работе, основан на модифицированной модели Хата, которая учитывает рельеф местности в местах нахождения базовой и портативной станций. Также алгоритм учитывает поправочный коэффициент на дополнительные потери, связанные с реальными условиями распространения радиоволн.

Задача автоматического размещения БС ТС может быть представлена в следующем виде:

– расчет допустимого уровня потерь на трассе распространения радиосигнала [9];

– расчет предварительного радиуса зоны обслуживания БС по модифицированной модели Хата;

– уточненный расчет зоны обслуживания БС с учетом потерь на дифракцию;

– вывод поправочного коэффициента радиуса зоны обслуживания, обеспечивающего качественную связь в любой точке нахождения ПС.

Алгоритм автоматического размещения базовых станций транкинговых систем связи

В статистической модели Хата не учитывается рельеф местности, что в реальности может привести к отсутствию сигнала в различных местах нахождения ПС при расчете зоны обслуживания. Для предварительного расчета радиуса зоны обслуживания предлагается модифицировать модель Хата и учитывать разницу в высотах над уровнем моря Δh в местах нахождения БС и ПС [10]. Напри-

мер, если БС находится в точке с высотой 212 м над уровнем моря, а ПС находится в точке с высотой 208 м над уровнем моря, то разницу в 4 м прибавляем к высоте подвеса антенны БС. Тогда модель Хата для городской местности примет вид (1), где: L_U – уровень потерь на трассе распространения радиосигнала, дБ; f – частота передачи сигнала, МГц; h_{BS} – высота подвеса антенны БС, м; h_{PS} – высота подвеса антенны ПС, м; Δh – разница в высотах над уровнем моря в местах нахождения БС и ПС, м; d – расстояние между объектами, км. Модель Хата для пригородной местности (L_{SU}) и открытых сред (L_0) представлена в виде (2) и (3) соответственно [11].

$$L_U = 69{,}55 + 26{,}16 \times \log_{10} f - 13{,}82 \times \log_{10}(h_{BS}; h_{BS} \pm \Delta h) - [0{,}8 + (1{,}1 \times \log_{10} f - 0{,}7) \times \times (h_{PS}; h_{PS} \pm \Delta h)] + [44{,}9 - 6{,}55 \times \log_{10}(h_{BS}; h_{BS} \pm \Delta h)] \times \log_{10} d \text{ [дБ]}, \quad (1)$$

$$L_{SU} = L_U - 2 \times \left(\log_{10} \frac{f}{28}\right)^2 - 5{,}4 \text{ [дБ]}, \quad (2)$$

$$L_0 = L_U - 4{,}78 \times (\log_{10} f)^2 + 18{,}33 \times \log_{10} f - 40{,}94 \text{ [дБ]}. \quad (3)$$

Возможна и обратная ситуация, когда ПС находится в точке с большей высотой над уровнем моря. В таком случае разницу предлагается суммировать с высотой подвеса антенны ПС, либо вычитать из высоты подвеса антенны БС. Ограничением для предложенной модификации являются допустимые диапазоны значений высот подвеса антенн БС и ПС, за которые выходить нельзя. Для БС это значение от 30 до 100 м, для ПС – от 1 до 10 м. Поэтому данный способ не подходит для применения в горной местности, где слишком большие перепады высот.

После получения предварительного радиуса зоны обслуживания БС необходимо рассчитать дополнительные потери на дифракцию, расчет которых проводится по рекомендации Международного Союза Электросвязи [12]. Потери за счет дифракции вычисляются путем сочетания метода, основанного на конструкции Буллингтона, и метода расчета дифракции над сферической Землей. Этот метод обеспечивает оценку дифракционных потерь для трасс всех типов, в том числе, трасс над морем; над территорией, удаленной от моря; или над побережьем, независимо от того, является ли трасса гладкой или пересеченной, трассой прямой видимости или загоризонтной трассой. Сначала определяется промежуточная точка профиля с наибольшим наклоном линии от передатчика к точке, м/км:

$$S_{tim} = \max\left[\frac{g_i + 500 \times C_e \times d_i \times (d - d_i) - h_{tc}}{d_i}\right], \quad (4)$$

где g_i – суммарная высота над уровнем моря и высота препятствия в i-й точке, м; d_i – расстояние от передатчика до i-й точки профиля, км; d – длина трассы, км; h_1 – суммарная высота антенны и высота над уровнем моря в месте нахождения передатчика, м; h_{tc} – max(g; h_1), (м); C_e – эффективная кривизна Земли, км$^{-1}$.

Далее рассчитывается наклон линии от передатчика к приемнику, предполагая трассу прямой видимости:

$$S_{tr} = \frac{h_{rc} - h_{tc}}{d} \text{ [м/км]}, \quad (5)$$

где h_n – суммарная высота антенны и высота над уровнем моря в месте нахождения приемника, м; h_{rc} – max (g; h_n), м.

Затем в зависимости от того, является ли дифракционная трасса трассой прямой видимости или загоризонтной трассой, определяются потери за счет дифракции на остром краю для точки Буллингтона:

$$L_{uc} = J(v_b) \text{ [дБ]}, \quad (6)$$

где функция безразмерного параметра v определяется выражением (7). Потери за счет дифракции Буллингтона на трассе определяются выражением (8). Затем рассчитываются потери за счет дифракции над сферической Землей, рассчитав наименьшую высоту просвета между трассой над криволинейной Землей и лучом между антеннами – выражение (9), где: a_p – эффективный радиус Земли, км.

$$J(v) = 6{,}9 + 20 \times \log_{10}(\sqrt{(v - 0{,}1)^2 + 1} + v - 0{,}1), \quad (7)$$

$$L_{bull} = L_{uc} + (1 - e^{\frac{-L_{uc}}{6}}) \times (10 + 0{,}02d) \text{ [дБ]}, \quad (8)$$

$$h_{se} = \frac{\left(h_{tesph} - 500 \times \frac{d_{se1}^2}{a_p}\right) \times d_{se2} + (h_{resph} - 500 \times \frac{d_{se2}^2}{a_p}) \times d_{se1}}{d} \text{ [м]}. \quad (9)$$

И требуемый просвет для нулевых потерь за счет дифракции:

$$h_{req} = 17{,}456 \times \sqrt{\frac{d_{se1} \times d_{se2} \times \lambda}{d}} \text{ [м]}, \quad (10)$$

где λ – длина волны, м.

После чего рассчитываются потери за счет дифракции над сферической Землей методом интерполяции:

$$L_{dsph} = \left(1 - \frac{h_{se}}{h_{req}}\right) \times L_{dft} \text{ [дБ]}, \quad (11)$$

где L_{dft} – потери за счет дифракции над сферической Землей, определяемые первым членом, дБ.

Потери за счет дифракции для общей трассы теперь определяются выражением:

$$L_d = L_{bulla} + \max\{L_{dsph} - L_{bulls}, 0\} \text{ [дБ]}. \quad (12)$$

Получим потери за счет дифракции L_{dp}, не превышаемые в течение p % времени, используя выражение:

$$L_{dp} = L_{d50} + (L_{d\beta} - L_{d50}) \times F_i \text{ [дБ]}, \quad (13)$$

где $L_{d50} = L_d = L_{d\beta}$, при проценте времени среднего года, в течение которого превышается рассчитанный уровень сигнала p = 50 %, дБ; F_i – коэффициент интерполяции.

Средние основные потери передачи, обусловленные дифракцией L_{bd50}, определяются выражением:

$$L_{bd50} = L_{bfs} + L_{d50} \text{ [дБ]}, \quad (14)$$

где основные потери в свободном пространстве определяются выражением:

$$L_{bfs} = 92{,}45 + 20 \times \log_{10} f + \\ + 20 \times \log_{10} d \text{ [дБ]}. \quad (15)$$

Основные потери передачи, обусловленные дифракцией, которые не превышаются в течение p % времени (процент времени среднего года, в течение которого превышается рассчитанный уровень сигнала), определяются выражением:

$$L_{bd} = L_{b0p} + L_{dp} \text{ [дБ]}, \quad (16)$$

где основные потери передачи по линии прямой видимости, которые не превышаются в течение p % времени:

$$L_{b0p} = L_{bfs} + E_{sp} \text{ [дБ]}. \quad (17)$$

Рассчитанные потери необходимо учитывать при расчете допустимого уровня потерь на трассе распространения радиосигнала. Исходя из полученной зоны обслуживания, можно определить гарантированную зону обслуживания в виде вписанной окружности, на границах которой устанавливается качественная связь (рисунок 2).

Далее для различных территорий выводится поправочный коэффициент, который затем учитывается при расчетах зон обслуживания, оставшихся БС многозоновой ТС. Формула выражения поправочного коэффициента такова:

$$K = \frac{(d_1 - d_2)}{d_1} \times 100\ \%, \quad (18)$$

где d_1 – диаметр предварительной зоны обслуживания, км; d_2 – диаметр гарантированной зоны обслуживания, км.

Поправочный коэффициент рассчитывается заранее для всех территорий страны и при расчетах вычитается из предварительной зоны обслуживания, рассчитанной статистическим методом. Алгоритм, представленный в данной статье, хорошо подходит для территорий, располагающихся на Восточно-Европейской равнине, Западно-Сибирской равнине, а также в Арктической зоне. Применение алгоритма в городской местности малоэффективно из-за наличия большого количества объектов, на которых установка БС запрещена.

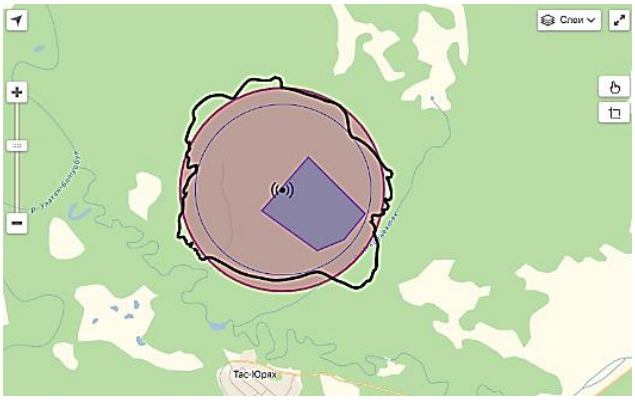


Рис. 2. Гарантированная зона обслуживания БС

Fig. 2. Guaranteed BS Service Area

Особенностью алгоритма является выполнение критерия минимизации, который заключается в минимальном количестве БС, необходимых для покрытия заданной территории связью.

На рисунке 3 представлен алгоритм автоматического размещения БС. На первом этапе происходит выбор территории на топографической карте, которую необходимо покрыть связью. Так как территория имеет сложную геометрическую форму, то необходимо определить минимальные и максимальные координаты по осям x и y и получить прямоугольный двумерный массив с определенной длиной ячейки (рисунок 4). Чем меньше ячейка, тем лучше будет результат, но тем больше будет и время расчета.

Далее происходит перебор всех точек, находящихся внутри выбранной территории. Опрошенная точка перемещается в базу посещенных. В конце каждой итерации происходит опрос на наличие непосещенных точек. Если таких точек не

осталось, то алгоритм завершает свою работу и выдает результат. Если остались непосещенные точки – алгоритм переходит к ним.

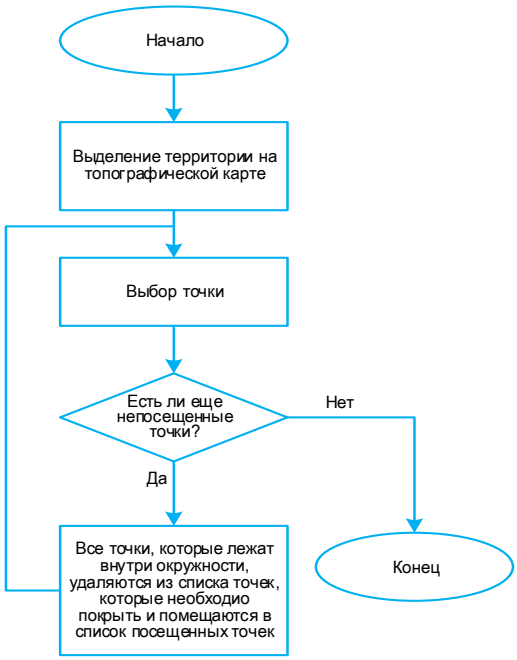


Рис. 3. Алгоритм автоматического размещения БС

Fig. 3. Algorithm for Automatic Placement of BS

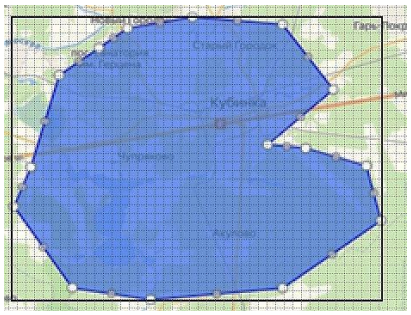


Рис. 4. Территория, которую необходимо покрыть связью

Fig. 4. The Territory that Needs to be Covered by the Connection

Выбор точки осуществляется также на основе алгоритма (рисунок 5). Определяются координаты точки $(x; y)$. Далее происходит опрос: является ли данная точка непосещенной? Если она таковой является, то для нее определяется эмпирически выведенная переменная s, которая зависит от радиуса зоны обслуживания БС (r) и количества точек (n):

$$s = r^3 + n^5, \qquad (19)$$

где r и n покрываются радиусом данной зоны обслуживания, но не покрыты другими радиусами.

Радиус зоны обслуживания определяется на основе расчетов по модифицированной модели Хата. Далее значение переменной s сравнивается с наилучшим значением этой переменной. Если это первая итерация, то наилучшее значение – 0. После чего запоминаются координаты данной точки, и алгоритм переходит к следующей точке. Так происходит до тех пор, пока не будут посещены все точки. После чего выводится точка, переменная s которой имеет наибольшее значение, и в этой точке строится окружность с рассчитанным ранее радиусом (рисунок 6). Далее происходит все то же самое, но в переборе уже не участвуют точки, которые находятся внутри окружности.

Рис. 5. Алгоритм выбора точки

Fig. 5. Point Selection Algorithm

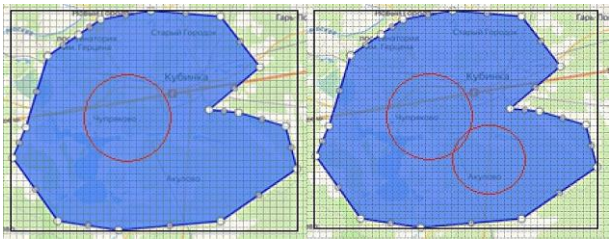


Рис. 6. Размещение БС в наилучших точках

Fig. 6. Placing the BS in the Best Points

Пример применения алгоритма

Рассмотрим работу алгоритма на примере размещения БС внутри территории, которую необходимо покрыть связью. Алгоритм применяется в веб-сайте, который на данный момент не доступен общему пользованию и находится в режиме тестирования. Для сравнения, полученные координаты внесем в программу RadioMobile для проверки полученных расчетов.

Пользователь на топографической карте местности выделяет территорию, которую необходимо покрыть связью, а также участки, в которых установка БС невозможна (водные препятствия, закрытые территории). Далее задается обязательный сокращенный набор параметров: среда распространения радиосигнала, частота передачи, высоты подвеса антенн.

Также существует расширенный набор параметров, а именно, технические характеристики приемопередающего оборудования: мощность передатчика, чувствительность приемника, коэффициенты усиления антенн, коэффициенты потерь в фидере. Если пользователь не обладает такой информацией, то он может оставить значение по умолчанию. После пользователю остается выбрать один из предложенных поправочных коэффициентов, который применяется для данной местности.

Для примера рассмотрим покрытие транкинговой связью территории возле п. Новый Бор Ямало-Ненецкого АО, где происходит нефтегазодобыча. Территории вокруг поселка полностью не покрыты мобильной связью операторов «Билайн» и «МТС», в связи с чем предприятию необходимо развернуть собственную связь (рисунок 7).

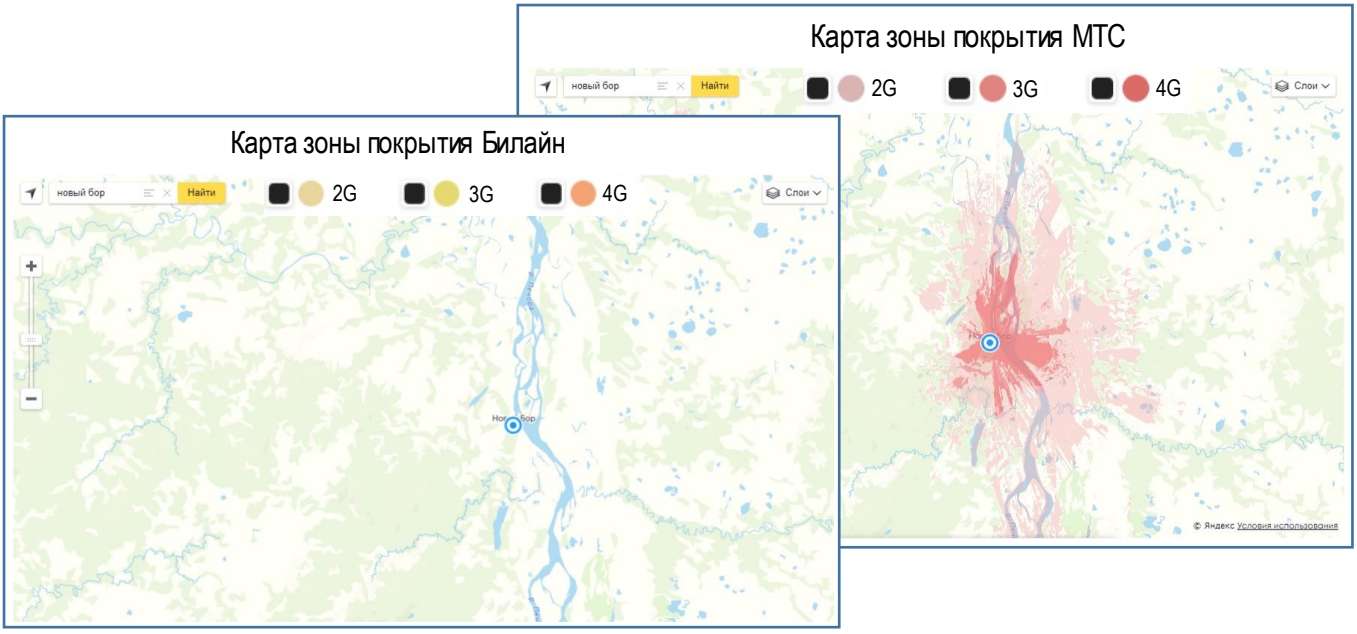


Рис. 7. Карта покрытия сотовых операторов «Билайн» (слева) и «МТС» (справа)

Fig. 7. Coverage Map of Beeline (on the Left) and MTS (on the Right) Mobile Operators

Синим цветом на карте выбрана территория, которую необходимо покрыть связью (рисунок 8). Алгоритм размещает 2 БС внутри замкнутого контура с выводом координат мест установки. Радиусы зон обслуживания БС составляют: R_1 = = 10000 м, R_2 = 10250 м. Затем полученные координаты БС1 (66.368312 С.Ш.; 51.754582 В.Д.) и БС2 (66.482650 С.Ш.; 51.659751 В.Д.) указываем в программе RadioMobile, задав при этом тот же самый набор параметров. В результате получаем те же самые радиусы зон гарантированного обслуживания, в которых нет зон с невозможностью приема. На рисунке 9 представлена зона обслуживания БС1, а на рисунке 10 – БС1 и БС2.

Рисунок 11 демонстрирует количество времени, затраченное на проектирование ТС связи в разработанном программном обеспечении (ПО) на основе предложенного алгоритма и проектирование в программе RadioMobile. Исходными данными, которые необходимо ввести, являются: мощность передатчика БС – 44дБм; мощность передатчика ПС – 30 дБм; чувствительность приемника БС – 106 дБм; чувствительность приемника ПС – 103 дБм; коэффициент потерь в фидере антенны БС – 6 дБм; коэффициент потерь в фидере антенны ПС – 0 дБм; коэффициент усиления антенны БС – 8 дБм; коэффициент усиления антенны ПС – 6 дБм; высота подвеса антенн БС – 50 м; высота подвеса антенн ПС – 1.5 м; частота передачи сигнала.

В разработанном ПО пользователем задаются все вышеперечисленные параметры, определяется территории, которые необходимо покрыть связью и внутри которых установка БС невозможна, и запускается расчет.

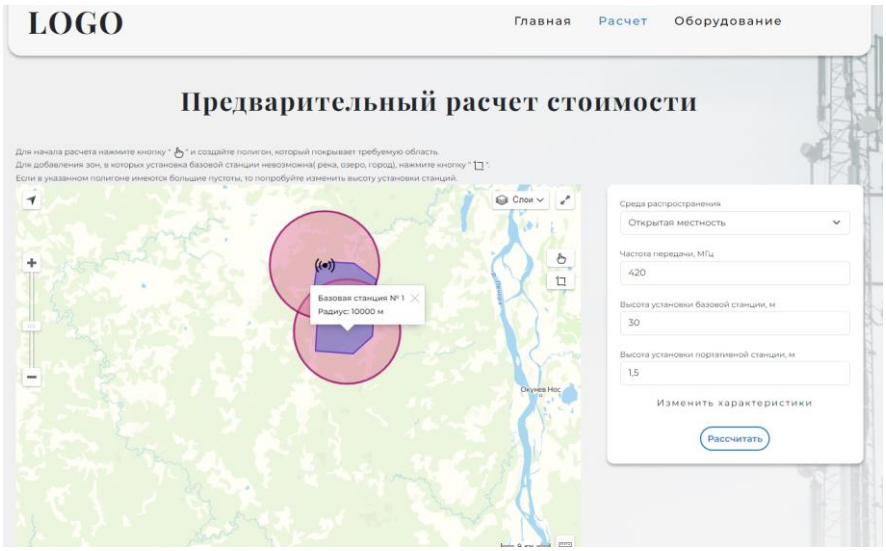


Рис. 8. Работа алгоритма

Fig. 8. Algorithm Operation

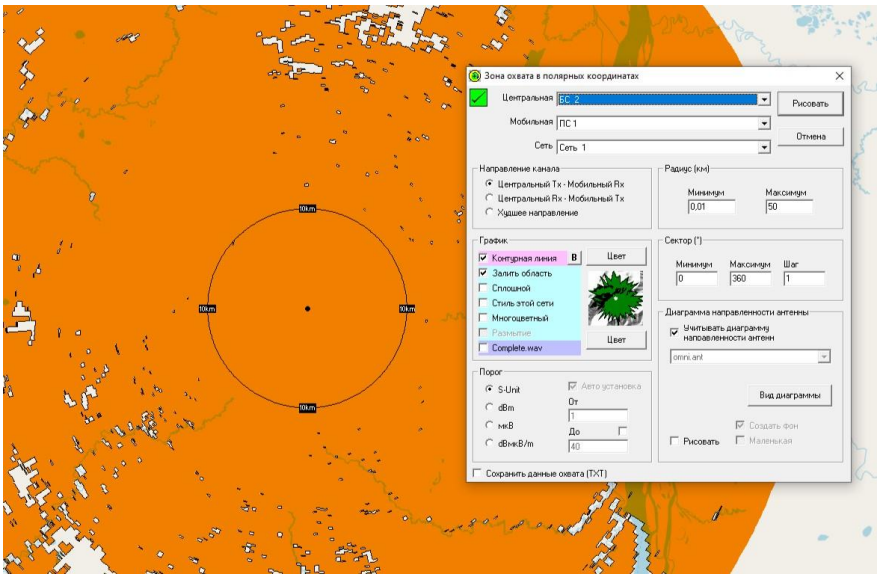


Рис. 9. Зона обслуживания БС1

Fig. 9. BS1 Service Area

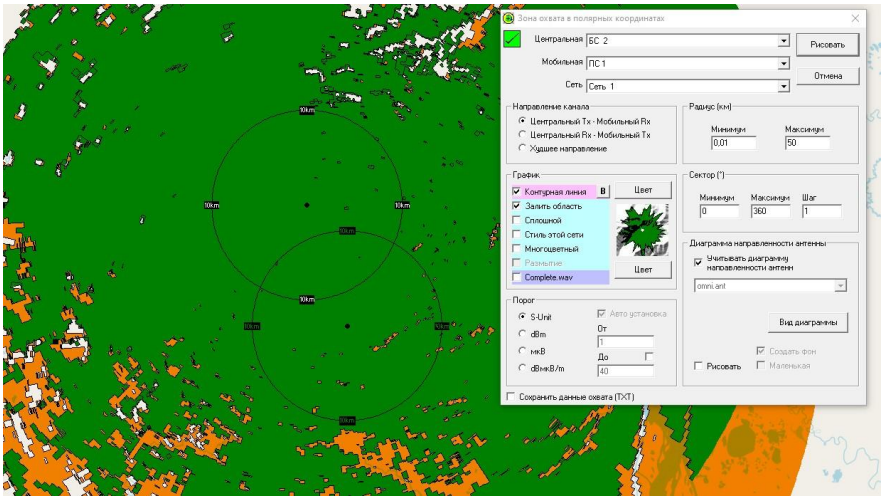


Рис. 10. Зона обслуживания БС1 и БС2

Fig. 10. BS1 and BS2 Service Area

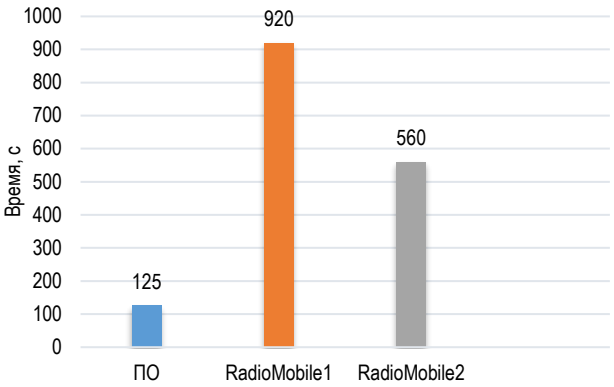


Рис. 11. Время, затраченное на проектирование ТС

Fig. 11. Time Spent on Design TS

В программе RadioMobile необходимо загрузить карту высот интересующего региона, а также указать ее размеры. Предполагается, что карта высот для всех регионов страны уже скачана на ПК пользователя и время на поиск и скачивание не учитывается. Далее для удобства отображения накладывается карта местности, указывается место размещения БС и задаются все ее вышеперечисленные параметры; затем строится карта зоны покрытия. В данной программе проведено два испытания: в первом случае расстановка БС осуществлялась вручную, на основе анализа топографической карты, во втором – БС размещались в местах, координаты которых получены из разработанного ПО.

Из результатов, представленных в гистограмме, следует, что количество времени, затраченное на проектирование ТС связи в разработанном ПО, в 4 раза меньше (125 с) времени, затраченного на проектирование в программе RadioMobile; при вводе координат мест установки БС (560 с), полученных из ПО, в 7 раз меньше, чем количество времени, затраченное на проектирование в программе RadioMobile при поиске мест размещения БС вручную (920 с).

Заключение

Представленный алгоритм автоматически размещает минимальное количество БС внутри территории, которую необходимо покрыть связью, с учетом статистического метода расчета и реальных условий распространения радиоволн.

Элементом новизны данного решения является то, что на первом этапе расчета определяется зона обслуживания БС статистическим методом, на втором этапе происходит уточненный расчет зоны обслуживания с учетом реальных условий распространения радиоволн на месте привязки, а на третьем этапе определяется зона гарантированного обслуживания абонентов. К исследованиям, в которых предлагается автоматическое размещение БС, можно отнести работы для сотовой связи [2–6, 8]. Однако в данных работах используются другие алгоритмы, в которых не учитываются реальные условия распространения радиоволн на месте нахождения БС и ПС. Проведено сравнение полученных результатов в программе RadioMobile, в ходе которого подтверждены зоны гарантированного обслуживания БС, а также получены данные, показывающие, что время проектирования ТС связи в разработанном ПО на основе представленного алгоритма в 4 раза меньше, чем время проектирования ТС в программе RadioMobile с известными координатами мест расположения БС.

Алгоритм, представленный в работе, может быть применен для проектирования ТС связи при освоении Арктической зоны, а также на равнинных территориях. Его использование для размещения БС в городе нецелесообразно, так как установка БС возможна лишь в определенных и разрешенных местах. В дальнейшем планируется доработка ПО и разработка метода размещения БС, методики проектирования зон обслуживания БС, основанных на представленном алгоритме.

Список источников

1. Сакалема Д.Ж. Подвижная радиосвязь. М.: Горячая линия – Телеком, 2012. 512 с.

2. Бабков В.Ю., Вознюк М.А., Михайлов П.А. Сети мобильной связи. Частотно-территориальное планирование. М.: Горячая линия – Телеком, 2007. 224 с.

3. Егоров Л.Л., Кологривов В.А., Мелихов С.В. Алгоритм расчета зон покрытия базовых станций сотовой связи // Доклады Томского Государственного Университета систем управления и радиоэлектроники. 2009. № 1-1(19). С. 15–19.

4. Мухаджинов Р.Р. Применение генетического алгоритма к решению задачи «Размещение станций систем мобильной связи» // Вестник Астраханского Государственного Технического Университета. Серия: управление, вычислительная техника и информатика. 2009. № 1. С. 165–167.

5. Павловская К.А. Применение генетического алгоритма для решения задач размещения базовых станций в сетях пятого поколения // Информатика и кибернетика. 2019. № 4(18). С. 29–34.

6. Мухтаров А.А., Першин О.Ю. Размещение базовых станций широкополосной беспроводной сети связи для обслуживания заданного множества рассредоточенных объектов // Двенадцатая международная конференция «Управление развитием крупномасштабных систем mlsd'2019» (Москва, Россия, 01–03 октября 2019). М.: Институт проблем управления им. В.А. Трапезникова РАН, 2019. С. 609–612. DOI:10.25728/mlsd.2019.1.0610

7. Аминова Р.Р. Разработка алгоритма первоначального размещения базовых станций сетей широкополосного радиодоступа на этапе частотно-территориального планирования // Всероссийская научно-практическая конференция с международным участием «Новые технологии, материалы и оборудование российской авиакосмической отрасли» (АКТО-2016, Казань, Россия, 10–12 августа 2016). Казань: Академия наук Республики Татарстан, 2016. Т. 2. С. 338–342.

8. Скаков Е.С., Малыш В.Н. Использование алгоритмов мультистарта и поиска с запретами для решения задачи размещения базовых станций // Информационно-управляющие системы. 2015. № 3(76). С. 99–106. DOI:10.15217/issn1684-8853.2015.3.99

9. Иванов В.С., Хаджийская Е.Ю. Расчёт зоны покрытия транкинговой системы связи // XIX Международная научно-практическая конференция «Инновационные, информационные и коммуникационные технологии» (Сочи, Россия, 01–10 октября 2022). М.: Ассоциация выпускников и сотрудников ВВИА имени профессора Н.Е. Жуковского содействия сохранению исторического и научного наследия ВВИА имени профессора Н.Е. Жуковского, 2022. С. 345–350.

10. Иванов В.С., Увайсов С.У., Иванов И.А. Алгоритм расчета зоны обслуживания базовой станции транкинговой системы связи // Наукоемкие технологии. 2023. Т. 24. № 4. С. 12–20. DOI:10.18127/j19998465-202304-02

11. Овчинников А.М., Воробьев С.В., Сергеев С.И. Открытые стандарты цифровой транкинговой радиосвязи. М: Эко-Трендз, 2000. 166 с.

12. Рекомендация МСЭ-R P.1812-4 (07/2015). Метод прогнозирования распространения сигнала на конкретной трассе для наземных служб "из пункта в зону" в диапазонах УВЧ и ОВЧ.

References

1. Sakalema D.J. *Mobile Radio Communication*. Moscow: Goriachaia liniia Telekom Publ.; 2012. 512 p.

2. Babkov V.Ju., Voznjuk M.A., Mihajlov P.A. *Mobile Communication Networks. Frequency-Territorial Planning*. Moscow: Goriachaia liniia – Telekom Publ.; 2007. 224 p.

3. Egorov L.L., Kologrivov V.A., Melihov S.V. Algorithm for Calculating Coverage Areas of Cellular Base Stations. *Proceedings of TUSUR University*. 2009;1-1(19):15–19.

4. Muhadzhinov R.R. Application of the genetic algorithm to the solution of the problem "Placement of stations of mobile communication systems. *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*. 2009;1:165–167.

5. Pavlovskaja K.A. A Genetic Algorithm is Used to Solve the Problems of Placing Base Stations in Fifth-Generation Networks. *Informatika i kibernetika*. 2019;4(18):29–34.

6. Muhtarov A.A., Pershin O.Ju. Placement of base stations of a broadband wireless communi-cation network to serve a given set of dispersed objects. *Proceedings of the XIIth International Conference on Management of Large-Scale Systems Development mlsd'2019", 01–03 October 2019, Moscow, Russia*. Moscow: V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences Publ.; 2019. p.609–612. DOI:10.25728/mlsd.2019.1.0610

7. Aminova R. Development of the Initial Base Stations Placement Algorithm for Broadband Radio Access Networks on the Frequency-Territorial Planning Stage. *Proceedings of the All-Russian Scientific and Practical Conference with International Participation on New Technologies, Materials and Equipment of the Russian Aerospace Industry, AKTO-2016, 10–12 August 2016, Kazan, Russia, vol.2*. Kazan: Academy of Sciences of the Republic of Tatarstan Publ.; 2016. p.338–342.

8. Skakov E.S., Malysh V.N. Multi-Start and Tabu Search Algorithms in Base Station Location Problem. *Information and Control Systems*. 2015;3(76):99–106. DOI:10.15217/issn1684-8853.2015.3.99

9. Ivanov V.S., Hadzhijskaja E.Ju. Calculation of coverage area of trunking communication system. *Proceedings of the XIXth International Scientific and Practical Conference on Innovation, Information and Communication Technologies, 01-10 October 2022, Sochi, Russia*. Moscow: Assotsiatsiia vypusknikov i sotrudnikov VVIA imeni professora N.E. ZHukovskogo sodeistviia sokhraneniiu istoricheskogo i nauchnogo naslediia VVIA imeni professora N.E. ZHukovskogo Publ.; 2022. p.345–350.

10. Ivanov V.S., Uvajsov S.U., Ivanov I.A. Algorithm for Calculating the Service Area of the Trunking Communication System Base Station. *Science Intensive Technologies*. 2023;24(4):12–20. DOI:10.18127/j19998465-202304-02

11. Ovchinnikov A.M., Vorobyov S.V., Sergeev S.I. *Open Standards of Digital Trunking Radio Communication*. Moscow: Eco-Trends Publ.; 2000. 166 p.

12. Rec. ITU-R P.1812-4. *A path-specific propagation prediction method for point-to-area terrestrial services in the UHF and VHF bands*. 2015.

Статья поступила в редакцию 14.07.2023; одобрена после рецензирования 14.10.2023; принята к публикации 16.10.2023.
The article was submitted 14.07.2023; approved after reviewing 14.10.2023; accepted for publication 16.10.2023.

Информация об авторах:

ИВАНОВ Вячеслав Сергеевич
аспирант кафедры конструирования и производства радиоэлектронных средств МИРЭА – Российский технологический университет
https://orcid.org/0000-0001-9827-1690

УВАЙСОВ Сайгид Увайсович
доктор технических наук, профессор, заведующий кафедрой конструирования и производства радиоэлектронных средств МИРЭА – Российский технологический университет
https://orcid.org/0000-0003-1943-6819

ИВАНОВ Илья Александрович
кандидат технических наук, доцент кафедры конструирования и производства радиоэлектронных средств МИРЭА – Российский технологический университет
https://orcid.org/0000-0003-1266-0228

Научная статья
УДК 004.8
DOI:10.31854/1813-324X-2023-9-5-35-42

Разработка и исследование системы автоматического распознавания цифр йеменского диалекта арабской речи с использованием нейронных сетей

Наим Хуссейн Али Радан ✉, naeem.radan@gmail.com
Константин Владимирович Сидоров, bmisidorov@mail.ru

Тверской государственный технический университет,
Тверь, 170026, Российская Федерация

Аннотация: *В статье описаны результаты исследований по разработке и тестированию системы автоматического распознавания речи (САРР) на арабских цифрах с помощью искусственных нейронных сетей. Для проведения исследований использовались звукозаписи (речевые сигналы) арабского йеменского диалекта, записанные в Республике Йемен. САРР представляет собой изолированную систему распознавания целых слов, она реализована в двух режимах: «дикторозависимая система» (дикторы при обучении и тестировании системы используются одни и те же) и «дикторонезависимая система» (дикторы, используемые для обучения системы, отличаются от тех, которые применяются для ее тестирования). В процессе распознавания речевой сигнал очищается от шумов с помощью фильтров, далее сигнал предварительно локализуется, обрабатывается и анализируется окном Хэмминга (применяется алгоритм временного выравнивания для компенсации различий в произношении). Информативные признаки извлекаются из речевого сигнала с использованием мел-частотных кепстральных коэффициентов. Разработанная САРР обеспечивает высокую точность распознавания арабских цифр йеменского диалекта – 96,2 % (для дикторозависимой системы) и 98,8 % (для дикторонезависимой системы).*

Ключевые слова: *нейронные сети, распознавание речи, йеменский диалект*

Ссылка для цитирования: Радан Н.Х.А., Сидоров К.В. Разработка и исследование системы автоматического распознавания цифр йеменского диалекта арабской речи с использованием нейронных сетей // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 35–42. DOI:10.31854/1813-324X-2023-9-5-35-42

Developed and Studied the Automatic Digit Recognition System for Yemeni Dialect of Arabic Using Neural Networks

Naeem Radan ✉, naeem.radan@gmail.com
Konstantin Sidorov, bmisidorov@mail.ru

Tver State Technical University,
Tver, 170026, Russian Federation

Abstract: *The article describes the results of research on the development and testing of an automatic speech recognition system (SAR) in Arabic numerals using artificial neural networks. Sound recordings (speech signals) of the Arabic Yemeni dialect recorded in the Republic of Yemen were used for the research. SAR is an isolated system of recognition of whole words, it is implemented in two modes: "speaker-dependent system" (the same speakers are used for training and testing the system) and "speaker-independent system" (the speakers used for training the sys-*

tem differ from those used for testing it). In the process of speech recognition, the speech signal is cleared of noise using filters, then the signal is pre-localized, processed and analyzed by the Hamming window (a time alignment algorithm is used to compensate for differences in pronunciation). Informative features are extracted from the speech signal using mel-frequency cepstral coefficients. The developed SAR provides high accuracy of the recognition of Arabic numerals of the Yemeni dialect – 96.2 % (for a speaker-dependent system) and 98.8 % (for a speaker-independent system).

Keywords: *neural networks, speech recognition, Yemeni dialect*

For citation: Radan N.H., Sidorov K. Development and Research of a System for Automatic Recognition of the Digits Yemeni Dialect of Arabic Speech Using Neural Networks. *Proceedings of Telecommun. Univ.* 2023;9(5):35–42. DOI:10.31854/1813-324X-2023-9-5-35-42

Введение

Современный стандартный арабский язык (MSA, *аббр. от англ.* Modern Standard Arabic) является семитским языком и на сегодняшний день является одним из древнейших языков в мире. В настоящее время MSA является пятым широко используемым языком в мире. MSA является первым языком в арабском мире, то есть в Саудовской Аравии, Иордании, Омане, Йемене, Египте, Сирии, Ливане и т. д. Арабские алфавиты используются в нескольких языках, таких как персидский, урду и малайский. MSA имеет в основном 34 фонемы, из которых шесть основных гласных и 28 согласных. Фонема представляет собой наименьший элемент речевой единицы, который указывает на различие в значении слова или предложении. В MSA меньше гласных, чем в английском. В нем три долгих и три кратких гласных, в то время как в американском английском не менее двенадцати гласных. Арабские фонемы состоят из двух различных классов, называемых фарингеальными и эмфатическими фонемами. Два класса встречаются только в семитских языках, таких как иврит, персидский и урду [1–4].

Особенности и характеристики произнесенных цифр

Задача автоматического распознавания произнесенных цифр является одной из самых сложных задач в области компьютерного распознавания речи. Процесс распознавания произнесенных цифр необходим во многих приложениях, требующих ввода цифр, таких как набор телефонных номеров с помощью речи, адресов, бронирование авиабилетов, автоматический справочник для приема или отправки информации и т. д. Арабский йеменский диалект (Республика Йемен) подвергся ограниченному количеству исследований по сравнению с другими языками, такими как английский, японский, русский и арабские диалекты других стран арабского мира.

На настоящий момент проведено несколько независимых исследований по распознаванию арабских цифр. В [5] разработана дикторонезависимая система автоматического распознавания речи (САРР) арабских цифр. Система разработана с использованием параметров LPC (*аббр. от англ.* Linear Predictive Coding) для выделения признаков и логарифмического отношения правдоподобия для измерения сходства. В [6] реализована САРР арабских цифр, которая достигла точности распознавания 97 %. Обе упомянутые выше системы являются изолированными системами распознавания слов. В [7] разработана САРР арабских гласных, дополнительно реализовано распознавание изолированных арабских гласных и изолированных арабских слов.

В рамках работы исследована силлабическая природа арабского языка с точки зрения типов слогов, структур слогов и основных правил написания ударения. Арабские цифры от нуля до девяти (Sifr, Waḥid, Ithniyn, Thalathah, Arbaah, Khamsih, Sittih, Sabaah, Thamaniyah, Tisaah) являются многосложными словами, за исключением первого, «нуля», который является односложным словом. Допустимые слоги в арабском языке: *CV*, *CVC* и *CVCC*, где *V* обозначает (долгую или короткую) гласную, а *C* – согласную. Арабские высказывания могут начинаться только с согласной [8]. В таблице 1 показаны десять арабских цифр – I, их арабское написание – II, фонетическое название – III, способ их произношения – IV, а также типы слогов – V и их количество – VI в каждой произносимой цифре.

ТАБЛИЦА 1. Арабские цифры

TABLE 1. Arabic Digits

I	II	III	IV	V	VI
0	صِفْر	Sifr	Sifr	*CVCC*	1
1	واحِد	Wahid	Wahid	*CV–CV*	2
2	إِثْنَيْن	ʔiθnajn	Ithnaiyn	*CVC–CVCC*	2
3	ثَلاثَة	θalɑːθih	Thalathah	*CV–CV–CVC*	3
4	أَرْبَعَة	ʔarbʕah	Arbaah	*CVC–CV–CVC*	3
5	خَمْسَة	xamsih	Khamsih	*CVC–CVC*	2
6	سِتَّة	Sittih	Sittih	*CVC–CVC*	2
7	سَبْعَة	sabʕah	Sabaah	*CVC–CVC*	2
8	ثَمانِيَة	θamaniːh	Thamaniyah	*CV–CV–CV–CVC*	4
9	تِسْعَة	tissʕah	Tisaah	*CVC–CVC*	2

Искусственные нейронные сети

Искусственные нейронные сети (ИНС) уже много лет применяются в области автоматического распознавания речи с целью достижения производительности сети, близкой к человеческой. Модели ИНС состоят из множества нелинейных вычислительных звеньев, работающих параллельно по схемам, аналогичным биологическим нейронным сетям [8]. ИНС широко использовались в области распознавания речи в течение последних трех десятилетий. Наиболее полезными характеристиками ИНС для решения задачи распознавания речи являются отказоустойчивость и свойство нелинейности [9].

Модели ИНС отличаются топологией сети, характеристиками узла и правилами обучения. Одной из важных моделей нейронных сетей являются многослойные персептроны (МП), которые представляют собой сеть прямой связи с нулем, одним или несколькими скрытыми слоями узлов между входными и выходными узлами [8]. Возможности МП происходят из-за нелинейностей, используемых с его узлами. Любая сеть МП должна состоять из одного входного слоя (не вычислительных, а исходных узлов), одного выходного слоя (вычислительных узлов) и нуля или более скрытых слоев (вычислительных узлов) в зависимости от сложности сети и требований приложения [9].

В данной статье описана система, автоматически распознающая арабские цифры (в йеменском диалекте). Для проведения исследований применены звукозаписи (РС – речевые сигналы) арабского йеменского диалекта, записанные в разных городах республики Йемен от нескольких дикторов мужского пола. Система разработана с использованием ИНС. Исследование проводилось в два этапа: на первом – разработана и исследована дикторозависимая система (т. е. при обучении и при тестировании системы использован один и тот же набор дикторов с разными произношениями цифр), а на втором этапе – исследована дикторонезависимая система (т. е. набор дикторов, используемых при обучении системы, отличается от набора дикторов, используемых при ее тестировании). Система разработана и исследована с использованием МП, сеть имеет три скрытых слоя. В качестве функции активации используется сигмоидальная функция (логическая – Logsig, линейная – Purelin).

Методика проведения экспериментов

Система автоматического распознавания речи (САРР) разделена на несколько модулей в соответствии с их функциональностью, как показано на рисунке 1. Входной модуль цифровой обработки сигналов, функции которого заключаются в получении речи через микрофон, фильтрации и дискретизации. Для фильтрации РС перед обработкой использован полосовой фильтр с частотами среза 100 Гц и 4,8 кГц. Частота дискретизации установлена на 16 кГц с 16-битным разрешением для всех записанных РС.

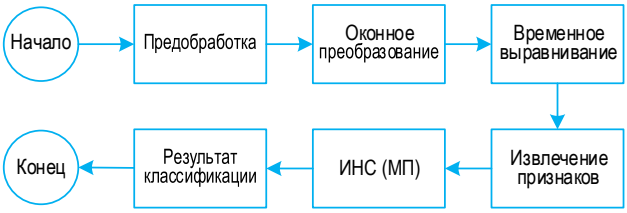


Рис. 1. Структурная схема проведения экспериментов

Fig. 1. Block Diagram of Experiments

Для отделения речи от отдельных частей сигнала, а также для определения начальной и конечной точек произносимого слова (цифры) использован метод ручного обнаружения (создан собственный алгоритм для выполнения текущей задачи). В каждом случае, чтобы выбрать точки данных для анализа РС, применено окно Хэмминга размером 256. В целях извлечения информативных признаков использованы мел-частотные кепстральные коэффициенты (МЧКК), для каждого сегмента извлекались 12 коэффициентов. При расчете МЧКК рассматривались 26 треугольных полосовых фильтров, структурная схема формирования МЧКК представлена на рисунке 2.

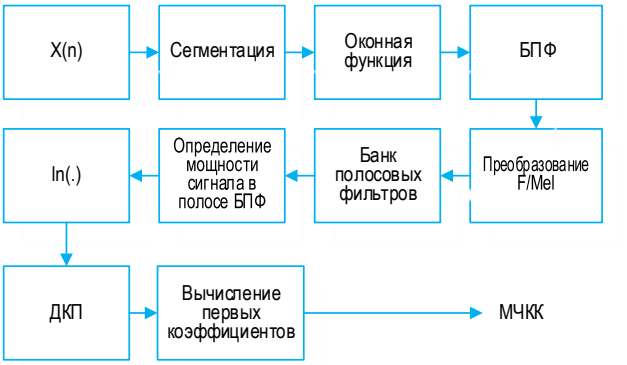


Рис. 2. Структурная схема формирования МЧКК

Fig. 2. Block Diagram of Procedure Extraction of Mel-Frequency Cepstral Coefficients

Сеть МП содержит три скрытых слоя со 150 нейронами в первом скрытом слое, с 75 – во втором и с 38 – в третьем скрытом слое. Выходной слой состоит из 10 нейронов. Каждый нейрон в выходном слое должен быть включен или выключен в зависимости от применяемой цифры во входном слое. Для нормальной и предполагаемой ситуации только один нейрон должен быть включен, в то время как все остальные – отключены, если применяемое высказывание является одной из десяти арабских цифр, в противном случае все нейроны должны быть отключены.

Извлечение информативных признаков

При формировании МЧКК рассматриваются следующие основные подходы:

1) звуковые колебания посредством микрофона преобразуются в РС;

2) после аналого-цифрового преобразования проводится сегментация РС;

3) каждый сегмент РС взвешивается оконной функцией;

4) взвешенные сегменты подвергаются быстрому преобразованию Фурье – формируется кратковременный спектр сигнала;

5) частотная шкала преобразуется в мел-шкалу (учитываются особенности человеческого слуха);

6) мел-частотный спектр сегмента равномерно разбивается на отдельные полосы набором полосовых фильтров;

7) определяется мощность сигнала на выходе каждого фильтра;

8) полученный набор значений мощностей сигналов логарифмируется;

9) к результату логарифмирования применяется дискретное косинусное преобразование (ДКП) – формируется кепстр РС.

Рассмотрим некоторые этапы алгоритма более подробно. МЧКК применяется в областях распознавания речи, при выделении признаков используется нелинейная шкала частот, представляющая собой шкалу Mel, для имитации частотной характеристики слуховой системы человека. МЧКК основаны на известном изменении критической полосы пропускания человеческого уха в зависимости от частоты. Также психоакустическая мера высоты тона, оцениваемая человеком, линейная в нижней части 1000 Гц и логарифмическая выше. МЧКК обеспечивают компактное представление данного РС. Математическая связь между шкалой частот Mel и линейной шкалой частот определяется следующим образом [10]:

$$f_{Mel} = 2595 + \log(1 + \frac{f_{HZ}}{700}), \quad (1)$$

где f_{HZ} – частота в Гц.

Предварительная обработка. Каждый сигнал, соответствующий каждой цифре, предварительно подчеркивается, чтобы увеличить отклик высоких частот РС: если $s(n)$ – исходный РС, а $s_p(n)$ – предварительно выделенный сигнал, то:

$$S_p(n) = S(n) - 0{,}97\, S(n-1) \quad (2)$$

подразумевает фильтрацию РС с использованием фильтра конечной импульсной характеристики, передаточная функция которого в области Z [11]:

$$h_p(z) = 1 - 0{,}97z^{-1}. \quad (3)$$

Оконное преобразование. Предварительно выделенный сигнал делится на кадры по 25 мс, т. е. для РС с частотой дискретизации, равной 16 кГц, получается, что длина кадра составляет 0,025 × 16000 = = 400 отсчетов, и умножается на перекрывающееся скользящее окно Хэмминга с шагом перекрытия 10 мс для подавления спектральных искажений в начале и в конце каждого кадра.

Окно Хэмминга рассчитывается по формуле, приведенной в [10]:

$$h(n) = \\ = \begin{Bmatrix} 0{,}45 - 0{,}46\cos\left(\frac{2\pi n}{N-1}\right), \text{если } 0 \le n \le N-1 \\ \text{или} \\ 0 \end{Bmatrix}, \quad (4)$$

где N – количество выборок в окне.

Дискретное преобразование Фурье. ДПФ используется для преобразования каждого кадра из N отсчетов из временной области в частотную, в результате получается спектр сигнала:

$$S(k) = \sum_{n=0}^{N-1} s(n) e^{-j2\pi kn/N}, k = 0,\ 1,\ 2,\ \ \ldots,\ N-1. \quad (5)$$

Банк полосовых фильтров. Поскольку диапазон частот, полученный на предыдущем шаге, широк, чтобы избежать вычислительных затрат, строится банк фильтров в шкале Mel. РС пропускается через банк, представляющий собой серию перекрывающихся треугольных фильтров, которые построены таким образом, что нижняя граница фильтра находится в центре предыдущего, а верхняя – в следующем фильтре. Предположим, что $Hm(k)$ – амплитудно-частотная характеристика m-го фильтра, где k – индекс дискретной частоты в цифровой области. Выход фильтра m-го фильтра Xm представляет мощность сигнала и может быть выражен как:

$$X_m = \sum_{k=0}^{\frac{N}{2}-1} |S(k)|^2\, |H_m(k)|, \quad 1 \le m \le k. \quad (6)$$

m – общее количество фильтров.

Дискретное косинусное преобразование. В результате применения ДКП (DCT, *аббр. от англ.* Discrete Cosine Transform) в сочетании с процедурой логарифмирования получится кепстр сигнала, представляющий МЧКК:

$$c(m) = DCT(\log(X_m). \quad (7)$$

Коэффициенты временной производной первого порядка МЧКК (ΔМЧКК), также известные как дифференциальные. Они соответствуют траекториям основных коэффициентов МЧКК и отражают их изменчивость во времени. ΔМЧКК рассчитываются по следующему уравнению регрессии [10]:

$$d_i = \frac{\sum_{n=1}^{N} n(c_{n+i} - c_{n-i})}{2\sum_{n+1}^{N} n^2}, \quad (8)$$

где d_i – дельта-коэффициент в кадре i, вычисленный с точки зрения соответствующих базовых кепстральных коэффициентов от c_{n+i} до c_{n-i}. Типичное значение N равно 2.

В результате использования данного подхода получаются компактные информативные признаки РС, а также сокращаются вычислительные и временные затраты при построении и исследовании системы распознавания речи [12].

Для распознавания неизвестной произнесенной цифры разработана сеть прямой связи в виде МП. При обучении сети МП использована логистическая нелинейная функция активации и алгоритм обратного распространения. Сеть состоит из N нейронов входного слоя. Их количество зависит от количества коэффициентов МЧКК для каждого кадра и количества рассматриваемых кадров РС, которые в данный момент подается на вход сети. Количество рассматриваемых кадров равно 111 в зависимости от используемого простого и эффективного алгоритма выравнивания по времени [7]: 12 коэффициентов МЧКК × 111 кадров = 1332.

База данных

Сформирована база данных, содержащая десять арабских цифр, полученная от 6 дикторов (носителей арабского йеменского диалекта) мужского пола. Объем базы данных состоит из 3 000 звукозаписей (РС), все дикторы произносили по 50 повторений для каждой цифры. Все звукозаписи от одного диктора записаны за один сеанс экспериментов. Все 3 000 звукозаписей (10 цифр ×50 повторений × 6 дикторов) использованы при обучении и тестировании САРР в зависимости от ее режима работы. Рассмотрены дикторозависимая и дикторонезависимая системы с параметрами: частота дискретизации – 16 кГц; база данных – 3000 звукозаписей; количество дикторов – 6; число повторений – 50; полосовой фильтр – 100 Гц и 4,8 кГц; оконная функция – Хэмминг; длительность сегмента – 256; коэффициент перекрытия – 128; функция активации – Logsig–Logsig–Logsig–Purelin; скрытые слои – 3; треугольные полосовые фильтры – 26.

Результаты и обсуждение

Дикторозависимая система

При исследовании дикторозависимой системы использованы произношения каждой цифры, которые произнесены всеми дикторами. Таким образом, общее количество звукозаписей, рассматриваемых для обучения, равно 1 500 звукозаписей (6 дикторов × 25 повторений × 10 цифр). При тестировании САРР использованы другие произношения каждой цифры из 1 500 звукозаписей. Таким образом, набор данных для обучения является подмножеством набора данных для тестирования. В таблице 2 (в ячейках слева от /) представлена матрица распознавания цифр, общая точность и ошибки данной системы.

ТАБЛИЦА 2. Матрица распознавания (дикторозависимая/дикторонезависимая система)

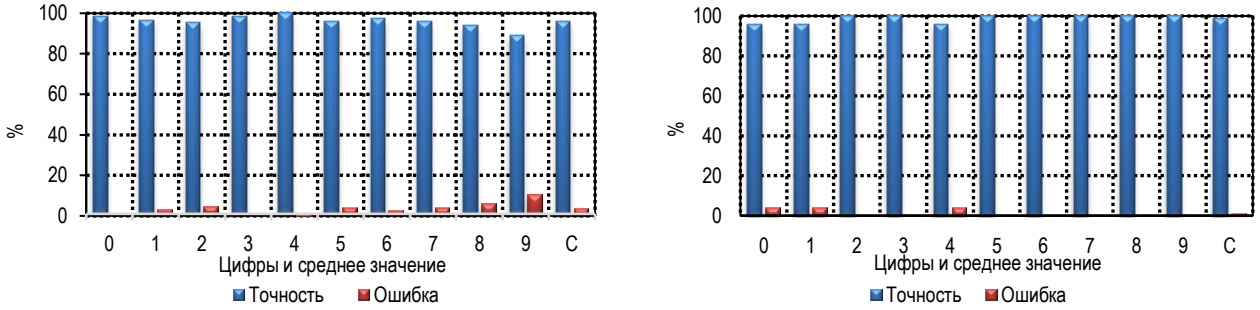
TABLE 2. Recognition Matrix (Speaker-Dependent / Speaker-Independent System)

Цифры	Ноль	Один	Два	Три	Четыре	Пять	Шесть	Семь	Восемь	Девять	Точность,%	Ошибки,%
Ноль	148 / 24	0 / 24	0 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	98,67 / 96,00	1,33 / 4,00
Один	0 / 0	145 / 1	4 / 25	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	12 / 0	96,67 / 96,00	3,33 / 4,00
Два	0 / 0	0 / 0	143 / 0	0 / 25	0 / 0	2 / 0	2 /	0 / 0	1 / 0	3 / 0	95,33 / 100,00	4,67 / 0,00
Три	0 / 1	4 / 0	0 / 0	148 / 0	0 / 24	1 / 0	0 / 0	0 / 0	1 / 0	0 / 0	98,67 / 100,00	1,33 / 0,00
Четыре	0 / 0	1 / 0	0 / 0	0 / 0	150 / 0	0 /25	0 / 0	4 / 0	0 / 0	1 / 0	100,00 / 96,00	0,00 / 4,00
Пять	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	144 / 0	2 / 0	0 / 0	7 / 0	0 / 0	96,00 / 100,00	4,00 / 0,00
Шесть	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	3 / 0	146 / 25	0 / 0	0 / 0	0 / 0	97,33 / 100,00	2,67/ 0,00
Семь	1 / 0	0 / 0	0 / 0	0 / 0	0 / 1	0 / 0	0 / 0	144 /25	0 / 0	0 / 0	96,00 / 100,00	4,00/ 0,00
Восемь	0 / 0	0 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	141 / 25	0 / 0	94,00 / 100,00	6,00/ 0,00
Девять	0 / 0	0 / 0	2 / 0	0 / 0	0 / 0	0 / 0	0 / 0	2 / 0	0 / 0	134	89,33 / 100,00	10,67/ 0,00
C											96,20 /98,80	3,80 / 1,20

В зависимости от набора тестовой базы данных система должна распознать 150 образцов для каждой цифры, где общее количество звукозаписей составляет 1 500 объектов. Общая средняя точность системы равна 96,20 %, что является достаточно высоким показателем, средняя ошибка системы составила 3,80 %. Системе не удалось распознать 57 объектов (0,038 × 1500) из 1 500 звукозаписей. Цифры 1, 2, 3, 4, 5, 6, 7 и 8 получили высокую точность распознавания. Наихудшая точность (89,33 %) получена при распознавании цифры 9. Несмотря на то, что размер базы данных невелик (всего десять произносимых арабских цифр), система продемонстрировала высокую производительность из-за вариативности произношения арабских цифр и того факта, что рассмотрен многоканальный режим в отличие от режима, зависящего от диктора, т. е. система обучается и тренируется

одними дикторами и разными произношениями. На рисунке 3 (слева) приведены зависимости точности и ошибки распознавания от конкретной цифры (от нуля до девяти). Также приведены средняя точность и средняя ошибка, которые обозначены буквой «*C*».

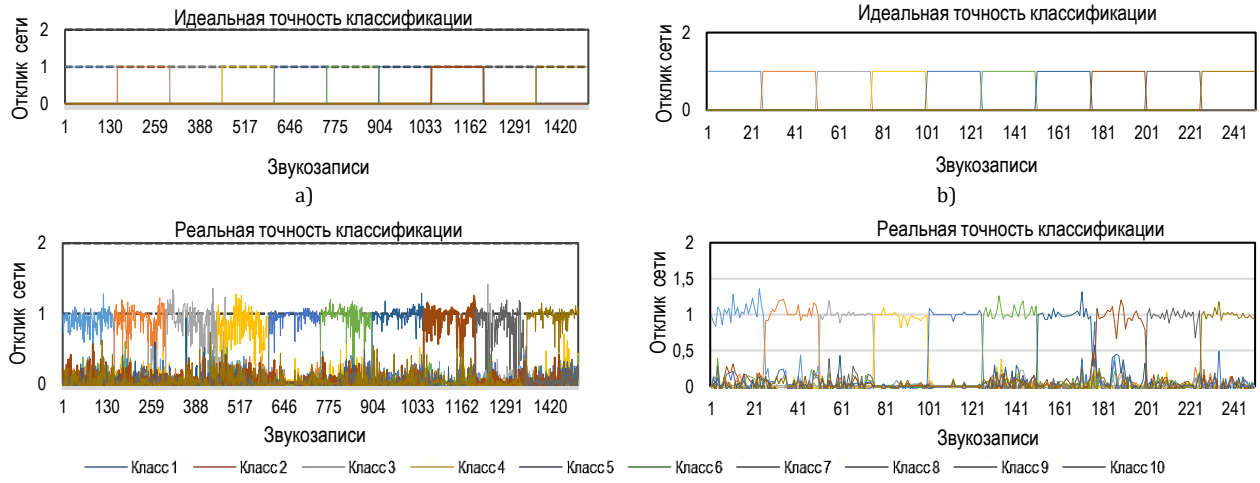
На рисунке 4 (слева) показаны идеальная и реальная точности классификации. Идеальная классификация проводится путем кодирования выхода сети 0 или 1, т. е. каждый нейрон в выходном слое должен быть включен или выключен в зависимости от применяемого значения во входном слое. При идеальной классификации для нормальной и предполагаемой ситуации только один нейрон должен быть включен, в то время как все остальные должны быть отключены, в случае реальной классификации выходы сети зависят от применяемого высказывания. Если им является одна из десяти арабских цифр, то тогда соответствующие нейроны должны быть включены. В противном случае все нейроны должны быть отключены, т. е. реальная классификация зависит от отклика сети и от сложности задачи распознавания.



a) b)

Рис. 3. Зависимость точности и ошибки распознавания от конкретной цифры для дикторозависимой (a) и дикторонезависимой (b) систем

Fig. 3. Recognition Accuracy Dependency on Specific Digits for Speaker-Dependent System (a) and Speaker-Independent System (b)



a) b)

c) d)

Рис. 4. Идеальная и реальная точность классификации САРР для дикторозавимисимой (слева) и дикторонезавимисимой (справа) систем

Fig. 4. Ideal and Real SAR Classification Accuracy for Speaker-Dependent System (Left) and Speaker-Independent System (Right)

Дикторонезависимая система

При исследовании дикторонезависимой системы использован один диктор для тестирования системы и шесть дикторов для обучения. Общее количество звукозаписей, предназначенных для тестирования, составляет 250 (1 диктор × 25 повторений × 10 цифр). Набор обучения состоит из 1 500 звукозаписей от шести дикторов. Все РС, подготовленные для обучения и тестирования САРР, представляют 1 750 звукозаписей (7 дикторов × 10 цифр × 25 повторений). В таблице 2 (в ячейках справа от /) показаны матрица распознавания цифр, общая точность и ошибки данной системы. Общее количество звукозаписей, протестированных САРР, составляет 250 объектов (1 диктор × 25 повторений × 10 цифр) для каждой цифры. Общая точность системы составляет 98,8 %, неправильно классифицированы 3 звукозаписи (0,012 × 250). Наихудшие результаты распознавания обнаружены в случаях с цифрами 0, 1 и 4, а наилучшие – с цифрами 2, 3, 5, 6, 7, 8 и 9.

На рисунке 3 (справа) представлены зависимости точности распознавания от конкретной цифры (от нуля до девяти), а также средняя точность и средняя ошибка *C*. На рисунке 4 (справа) продемонстрированы идеальная и реальная классификации. Классификация проводилась по процедуре, аналогичной классификации для дикторозависимой системы.

Таким образом, точность цифры 9 для дикторонезависимой системы составляет 100 %, для дикторозависимой системы – 89,33 %. Путем проверки спектрограмм цифр, форм сигналов, типов и количества слогов было обнаружено, что арабская цифра 9 акустически отличается от остальных цифр. Дополнительно обнаружено, что цифра 1 имеет высокий уровень ошибок. Ее анализ спектрограммы позволяет говорить о том, что есть сходство между цифрой 1 и цифрами 3, 2 и 9. В таблице 3 приведена вспомогательная информацию по сходству цифр.

ТАБЛИЦА 3. Некорректно классифицированные звукозаписи

TABLE 3. Incorrectly Classified Digits

Цифра	Некорректно классифицированные звукозаписи	
	Дикторонезависимая система	Дикторозависимая система
0	3	–
1	2, 3, 9	–
2	5, 6. 8, 9	1
3	1, 5, 8,	0
4	1, 7, 9	–
5	0, 6, 8	–
6	5	–
7	0	4
8	2	–
9	3	–

На рисунке 5 проиллюстрирован сравнительный анализ работы САРР в двух режимах. Точность дикторонезависимой системы превышает точность дикторозависимой системы. Конечно, дикторонезависимая система более практична, но для построения и разработки таких систем потребуется большой объем базы данных. Следует особо отметить тот факт, что полученные результаты не являются окончательными из-за ограничений в наборе исходных данных.

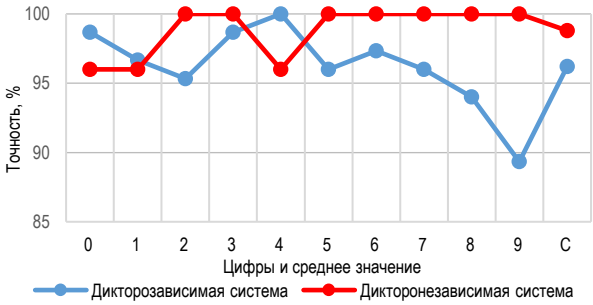


Рис. 5. Сравнительная точность САРР

Fig. 5. Comparative Accuracy of SAR System

Заключение

В рамках данной работы предложена система автоматического распознавания речи, протестированная с использованием звукозаписей цифр йеменского диалекта арабской речи. Система разработана с применением многослойных персептронов. При извлечении информативных признаков, с целью сжатии объема входных данных и сокращения времени работы системы, применен математический аппарат мел-частотных кепстральных коэффициентов. САРР работает в режиме дикторонезависимой и дикторозависимой систем. База данных объемом 3 000 звукозаписей создана с использованием 6 дикторов, которые являются носителями арабского йеменского диалекта. Общая точность работы САРР составляет 96,2 % (для дикторозависимой системы) и 98,8 % (для дикторонезависимой системы). На текущий момент времени авторы использовали для тестирования САРР свою небольшую базу арабских цифр, так как отсутствует проверенная и стандартная большая база цифр йеменского диалекта. В дальнейшем авторы планируют заняться задачами адаптации и оптимизации параметров САРР в шумных условиях, приближенных к реальным, а также увеличением объема базы данных. Дополнительно планируется сотрудничество с автором работы [11], в рамках которого для проверки работы, предложенной САРР, будут использованы записи телефонной речи в Республике Йемен.

Список источников

1. Al-Zabibi M. An acoustic-phonetic approach in automatic Arabic speech recognition. Loughborough University. Doctoral Thesis. 1990. URL: https://hdl.handle.net/2134/6949 (Accessed 02.10.2023)

2. Alkhouli M. Alaswaat Alaghawaiyah // Daar Alfalah, Jordan. 1990 (in Arabic)

3. Deller J., Hansen J., Proakis J. Discrete-Time Processing of Speech Signal. 1993. DOI:10.1109/9780470544402

4. Elshafei M. Toward an Arabic Text-to-Speech System // The Arabian Journal for Scince and Engineering. 1991. Vol. 16. Iss. 4B. PP. 565–583.

5. Hagos E. Implementation of an Isolated Word Recognition System. M.Sc. Thesis. King Fahd University of Petroleum & Minerals Dhahran, Saudi Arabia. 1985.

6. Abdulla W.H., Abdul-Karim M.A.H. Real-time spoken Arabic digit recognizer // International Journal of Electronics. 1985. Vol. 59. Iss. 5. PP. 645–648. DOI:10.1080/00207218508920741

7. Alotaibi Y.A. Investigating spoken Arabic digits in speech recognition setting. *Information Sciences*. 2005. Vol. 173. Iss. 1-3. PP. 115–139. DOI:10.1016/j.ins.2004.07.008

8. Alotaibi Y.A. High performance Arabic digits recognizer using neural networks // Proceedings of the International Joint Conference on Neural Networks (Portland, USA, 20–24 July 2003). IEEE, 2003. DOI:10.1109/ijcnn.2003.1223444

9. Alotaibi Y.A. Analyzing Arabic digit recognizer errors using spectrograms // Proceedings 7th International Conference on Signal Processing (ICSP, Beijing, China, 31 August 2004 – 04 September 2004). IEEE, 2004. DOI:10.1109/icosp.2004.1452746

10. Hassine M., Boussaid L., Massaoud H. Tunisian Dialect Recognition Based on Hybrid Techniques // International Arab Journal of Information Technology. 2018. Vol. 15. No. 1. PP. 58–65.

11. Аль-Дайбани А.М.С. Исследование методов и разработка алгоритмов обработки сигналов для систем автоматического распознавания телефонной речи в республике Йемен. Дис. … канд. техн. наук. Владимир: Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых, 2019. 150 с.

12. Радан Н.Х. Системы автоматического распознавания арабской речи и йеменского диалекта // Научно-аналитический журнал «Вестник Санкт-Петербургского университета государственной противопожарной службы МЧС России». 2023. № 2. С. 194–212.

References

1. Al-Zabibi M. *An acoustic-phonetic approach in automatic Arabic speech recognition. Loughborough University*. Doctoral Thesis. 1990. URL: https://hdl.handle.net/2134/6949 [Accessed 02.10.2023]

2. Alkhouli M. Alaswaat Alaghawaiyah. *Daar Alfalah, Jordan.* 1990 (in Arabic)

3. Deller J., Hansen J., Proakis J. *Discrete-Time Processing of Speech Signal*. 1993. DOI:10.1109/9780470544402

4. Elshafei M. Toward an Arabic Text-to-Speech System. *The Arabian Journal for Science and Engineering*. 1991;16(4B): 565–583.

5. Hagos E. *Implementation of an Isolated Word Recognition System*. M.Sc. Thesis. King Fahd University of Petroleum & Minerals Dhahran, Saudi Arabia. 1985.

6. Abdulla W.H., Abdul-Karim M.A.H. Real-time spoken Arabic digit recognizer. *International Journal of Electronics.* 1985; 59(5):645–648. DOI:10.1080/00207218508920741

7. Alotaibi Y.A. Investigating spoken Arabic digits in speech recognition setting. *Information Sciences*. 2005;173(1-3):115–139. DOI:10.1016/j.ins.2004.07.008

8. Alotaibi Y.A. High performance Arabic digits recognizer using neural networks. *Proceedings of the International Joint Conference on Neural Networks, 20–24 July 2003, Portland, USA*. IEEE; 2003. DOI:10.1109/ijcnn.2003.1223444

9. Alotaibi Y.A. Analyzing Arabic digit recognizer errors using spectrograms // Proceedings 7th International Conference on Signal Processing, ICSP, 31 August 2004 – 04 September 2004, Beijing, China. IEEE; 2004. DOI:10.1109/icosp.2004.1452746

10. Hassine M., Boussaid L., Massaoud H. Tunisian Dialect Recognition Based on Hybrid Techniques. *International Arab Journal of Information Technology*. 2018;15(1):58–65.

11. Al-Daibani A.M.S. *Research of methods and development of algorithms of signal processing for systems of automatic recognition of telephone speech in the Republic of Yemen*. PhD Thesis. Vladimir: Vladimir State University named after Alexander Grigorievich and Nikolai Grigorievich Stoletov Publ.; 2019. 150 p.

12. Radan N. Automatic speech recognition systems for Arabic speech and Yemeni dialect. Bulletin of St. Petersburg University of the State Fire Service of the Ministry of Emergency Situations of Russia. 2023;2:194–212.

Статья поступила в редакцию 03.03.2023; одобрена после рецензирования 28.03.2023; принята к публикации 26.09.2023.
The article was submitted 03.03.2023; approved after reviewing 28.03.2023; accepted for publication 26.09.2023.

Информация об авторах:

РАДАН Наим Хуссейн Али
аспирант кафедры информационных систем Тверского государственного технического университета
https://orcid.org/0009-0006-1723-2782

СИДОРОВ Константин Владимирович
кандидат технических наук, доцент, доцент кафедры автоматизации технологических процессов Тверского государственного технического университета
https://orcid.org/0000-0003-1119-2610

Научная статья
УДК 621.396.677
DOI:10.31854/1813-324X-2023-9-5-43-64

Диаграммообразование на основе позиционирования в сверхплотных сетях радиодоступа миллиметрового диапазона. Часть 2. Модель совокупности радиолиний

Григорий Алексеевич Фокин, fokin.ga@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

Аннотация: *Вторая часть исследования вопросов диаграммообразования на основе позиционирования в сверхплотных сетях радиодоступа диапазона миллиметровых волн посвящена формализации и программной реализации комплексной имитационной модели функционирования совокупности направленных радиолиний. Каждая направленная радиолиния между базовой станцией gNB (gNodeB), оборудованной антенной решеткой, и пользовательским устройством (UE, аббр. от англ. User Equipment), работающим в ненаправленном режиме, формируется по известному на gNB местоположению UE. Совокупность одновременно функционирующих в общем диапазоне частот направленных радиолиний gNB→UE исследуется как набор трафиковых лучей, реализующих множественный доступ с пространственным мультиплексированием (SDMA, аббр. от англ. Space-Division Multiple Access). Пространственное уплотнение реализуется посредством трехмерного диаграммообразования на базовой станции и позволяет компенсировать потери распространения радиоволн и высокий уровень помех. В первой части исследования было показано, что проблемой практической реализации SDMA в сверхплотных сетях радиодоступа является существенный (десятки дБ) разброс отношения сигнал/(шум + помеха) SINR (аббр. от англ. Signal Interference + Noise Ratio) в зависимости от взаимного расположения двух устройств. Целью настоящего исследования является установление зависимости SINR от 1) ширины луча сектора базовой станции gNB в направлении на пользовательское устройство UE в радиолинии полезного сигнала (SOI, аббр. от англ. Signal Of Interest); 2) неопределенности местоположения UE; 3) помех от радиолиний (SNOI, аббр. от англ. Signal Not of Interest): а) внутри своего сектора, б) других секторов своей соты и в) других сот сети. Разработанная и программно реализованная в настоящей работе имитационная модель впервые позволила установить взаимозависимость факторов погрешности позиционирования UE и требуемой ширины трафикового луча для его обслуживания. В частности, установлено, что с уменьшением погрешности позиционирования с 10 до 1 м требуемая ширина луча в горизонтальной и вертикальной плоскости сужается до 3 °, что позволяет увеличить SINR до 25 дБ. Исследование уплотнения одновременных передач показало, что для 64 пространственно мультиплексируемых UE с увеличением размера соты с 20 до 300 м отношение SINR увеличивается примерно на 30 дБ при ограничении на ширину луча в 3°. В отличие от похожих исследований в настоящей модели вклад от помех одновременно работающих трафиковых лучей внутри своего сектора, других секторов своей соты и других сот сети впервые показан по отдельности, что позволяет дифференцировать происхождение помех и использовать научно-обоснованное управление шириной луча для их компенсации.*

Ключевые слова: *диаграммообразование, позиционирование, сверхплотная сеть радиодоступа, миллиметровые волны, направленные радиолинии, ширина и ориентация луча, отношение сигнал/(шум + помеха)*

Источник финансирования: Исследование выполнено при финансовой поддержке Российского научного фонда (грант № 22-29-00528). https://rscf.ru/project/22-29-00528.

Ссылка для цитирования: Фокин Г.А. Диаграммообразование на основе позиционирования в сверхплотных сетях радиодоступа миллиметрового диапазона. Часть 2. Модель совокупности радиолиний // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 43–64. DOI:10.31854/1813-324X-2023-9-5-43-64

Location Aware Beamforming in Millimeter-Wave Band Ultra-Dense Radio Access Networks. Part 2. Model of a Set of Radio Links

Grigoriy Fokin, fokin.ga@sut.ru

The Bonch-Bruevich Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Abstract: *The second part of the study of beamforming issues, based on positioning in ultra-dense millimeter wave radio access networks, is devoted to the formalization and software implementation of a complex simulation model of the functioning of a set of directional radio links. Each directional radio link between a base station (gNodeB – gNB), equipped with an antenna array, and a user equipment (UE), operating in omnidirectional mode, is formed according to the location of the UE, known at the gNB. The set of gNB→UE directional radio links, simultaneously operating in a common frequency range, is studied as a set of traffic beams, that implement space division multiple access (SDMA). Spatial multiplexing is implemented through three-dimensional beamforming at the gNB and makes it possible to compensate for propagation losses and high levels of interference. In the first part of the study, it was shown that the problem of practical implementation of SDMA in ultra-dense radio access networks is a significant (tens of dB) spread in the signal to interference plus noise ratio (SINR), depending on the arrangement of two devices. The purpose of this study is to establish the dependence of SINR on 1) the beamwidth of the gNB sector in the direction of the user equipment in the radio link of the signal of interest (SOI); 2) uncertainty of the UE location; 3) interference from radio links of signal not of interest (SNOI): a) within its sector, b) other sectors of its cell and c) other cells in the network. The simulation model developed and implemented in software in this work for the first time made it possible to establish the interdependence of the UE positioning error factors and the required width of the traffic beam for its service. In particular, it was found, that as the positioning error decreases from 10 to 1 m, the required beam width in the horizontal and vertical planes narrows to 3 °, which makes it possible to increase the SINR to 25 dB. A simultaneous transmission multiplexing study showed that for 64 spatially multiplexed UEs, as the cell size increases from 20 to 300 m, the SINR increases by approximately 30 dB, subject to a beamwidth constraint of 3°. Unlike similar studies, in this model, the contribution from interference from simultaneously operating traffic beams within its sector, other sectors of its cell and other cells in the network is shown separately for the first time, which allows to differentiate the origin of interference and use scientifically based beamwidth control for their compensation.*

Keywords: *beamforming, positioning, ultra-dense radio access network, millimeter wave, directional radio links, beam width and orientation, signal interference + noise ratio*

Funding: the work was supported by the Russian Science Foundation, grant No. 22-29-00528, https://rscf.ru/project/22-29-00528

For citation: Fokin G. Location Aware Beamforming in Millimeter-Wave Band Ultra-Dense Radio Access Networks. Part 2. Model of a Set of Radio Links. *Proceedings of Telecommun. Univ.* 2023;9(5):43–64. DOI:10.31854/1813-324X-2023-9-5-43-64

1. ВВЕДЕНИЕ

Настоящее исследование является обобщением работы [1] на случай совокупности радиолиний. Концепция диаграммообразования на основе позиционирования в сверхплотных сетях радиодоступа (СРД) диапазона миллиметровых волн (ММВ) описана в [2] и основана на сценариях сетевого позиционирования пользовательских устройств [3]. Моделирование совокупности направленных радиолиний, выполненное ранее [4–6], не учитывало факторов погрешности позиционирования пользовательского устройства (UE, *аббр. от англ.* User Equipment) при ориентации луча базовой станцией gNB (gNodeB) в направлении на UE, что, однако существенно влияет на бюджет направленных радиолиний [7–9].

Диаграммообразование на основе позиционирования (LAB, *аббр. от англ.* Location Aware Beamforming) обозначено как один из основополагающих инструментов пространственного уплотнения одновременных передач (SDMA, *аббр. от англ.* (Space-Division Multiple Access) в сверхплотных СРД диапазона ММВ, который позволит компенсировать рост помех при увеличении плотности одновременно работающих устройств в общем диапазоне частот [10–15].

Анализ направленных радиолиний на уровне СРД ранее уже проводился [16–18], однако факторы погрешности позиционирования UE и ширины луча сектора базовой станции gNB при его ориентации в направлении на UE комплексно исследованы еще не были. Гипотезой настоящего исследования является утверждение о том, что чем точнее известны координаты UE, тем уже по азимуту и углу места может быть луч, сформированный сектором базовой станции gNB при диаграммообразовании на основе позиционирования для данного UE, при этом такой луч в условиях функционирования сверхплотной СРД может не перекрываться или лишь частично перекрываться с другими лучами на соседние UE, приводя к уменьшению отношения сигнал/(шум + помеха) SINR (*аббр. от англ.* Signal Interference + Noise Ratio).

Объектом исследования является совокупность радиолиний с диаграммообразованием на основе позиционирования в составе нескольких сот сверхплотной СРД диапазона ММВ.

Предметом исследования является зависимость отношения сигнал/(шум + помеха), а также требуемой ширины луча от точности позиционирования UE, размера соты и числа устройств в секторе.

Методом исследования является имитационное моделирование взаимного влияния совокупности направленных радиолиний внутри своего сектора, других секторов своей соты и других сот сети.

Целью исследования является установление влияния ориентации и ширины луча базовой станции, а также погрешности определения местоположения UE на уровень пространственного уплотнения одновременных передач по критерию отношения сигнал/(шум + помеха).

Материал исследования организован далее следующим образом. В разделе 2 выполнена формализация модели совокупности радиолиний с диаграммообразованием на основе позиционирования, а раздел 3 содержит ее программную реализацию. Результаты имитационного моделирования по оценке влияния ориентации и ширины луча, а также погрешности определения местоположения на уровень пространственного уплотнения одновременных передач по критерию отношения сигнал/(шум + помеха) представлены в разделе 4. Выводы сформулированы в разделе 5.

2. ФОРМАЛИЗАЦИЯ МОДЕЛИ СОВОКУПНОСТИ РАДИОЛИНИЙ С LAB

2.1. Постановка задачи моделирования

Формализуем задачу моделирования совокупности направленных радиолиний с диаграммообразованием на базовой станции gNB (gNodeB) на основе предварительного позиционирования UE. Рисунок 1 иллюстрирует сценарий модели диаграммообразования на основе позиционирования, учитывающий взаимное влияние направленных радиолиний. Метрикой оценки взаимного влияния направленных радиолиний является SINR.

На рисунке 1 показаны 7 сот, каждая из которых образована тремя секторами.

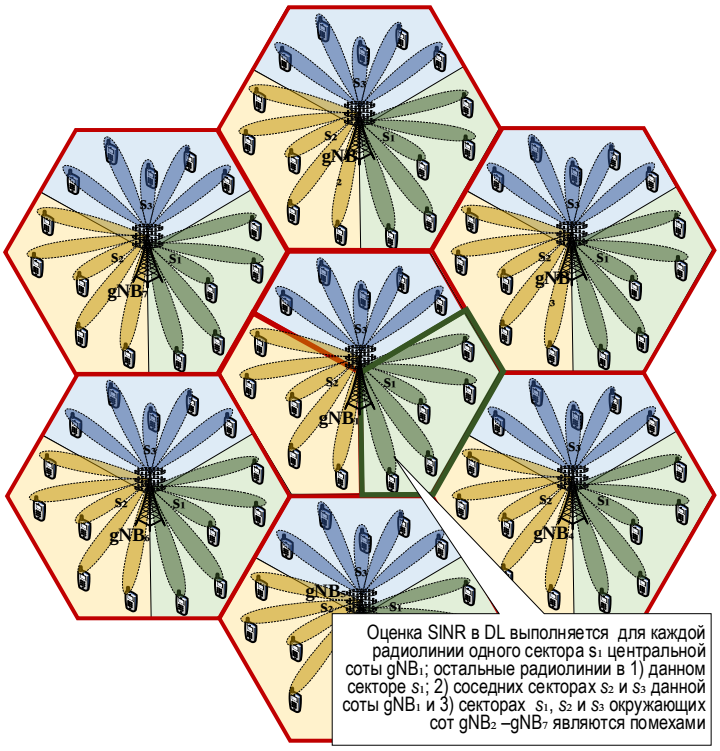


Рис. 1. Сценарий модели диаграммообразования на основе позиционирования

Fig. 1. Location-Aware Beamforming Model Operation Scenario

Каждую соту обслуживает базовая станция gNB_j, $j = 1, \dots, 7$. Каждый сектор s_i, $i = 1, 2, 3$ трехсекторной соты j оборудован многоканальным приемопередатчиком с антенной решеткой (АР), установленной на антенно-мачтовом устройстве базовой станции gNB_j. Возможности АР сектора базовой станции gNB позволяют одновременно формировать несколько лучей при работе на передачу и обслуживать несколько UE с организацией направленных радиолиний в нисходящем канале (DL, *аббр. от англ.* DownLink) от gNB к UE.

Допустим, что UE работают на прием в ненаправленном режиме. Пусть в каждом секторе s_i одновременно обслуживаются K пользовательских устройств UE_k, $k = 1, \dots, K$. Тогда в модели сети для обслуживания K UE в каждом секторе s_i базовой станции gNB_j необходимо организовать K направленных радиолиний $gNB_{js_i} \rightarrow UE_k$. Общее число одновременно работающих направленных радиолиний в представленной на рисунке 1 модели сети из семи трёхсекторных базовых станций равно $21K$. Ориентация диаграмм направленности АР секторов s_i базовых станций gNB_j осуществляется по известному местоположению пользовательских устройств UE_k в каждом секторе s_i.

Оценка SINR выполняется в DL для каждой направленной радиолинии $gNB_{1s_i} \rightarrow UE_k$ каждого сектора s_i центральной соты базовой станции gNB_1; границы одного из секторов на рисунке 1 выделены зеленым цветом. Радиолинии в данном секторе при оценке SINR являются полезными сигналами (SOI, *аббр. от англ.* Signal of Interest). Одновременно с радиолиниями SOI в модели сети на рисунке 1 работают радиолинии SNOI (*аббр. от англ.* Signal Not of Interest), которые при оценке SINR являются помехами. Для каждой k-й радиолинии SOI $gNB_{1s_i} \rightarrow UE_k$ в секторе s_i центральной соты базовой станции gNB_1 помехи образуются радиолиниями SNOI:

1) $gNB_{1s_i} \rightarrow UE_{k'}, k' \neq k$ внутри своего сектора s_i центральной соты базовой станции gNB_1;

2) других секторов своей соты $gNB_{1s_i}, \rightarrow UE_k, i' \neq i$ центральной базовой станции gNB_1;

3) $gNB_{js_i} \rightarrow UE_k$ секторов s_i окружающих сот базовых станций $gNB_j, j = 2, \dots, 7$.

При имитационном моделировании (ИМ) ненаправленных радиолиний в СРД диапазона дециметровых волн (ДМВ) учет помех от сот первого круга – 6-ти других сот, окружающих центральную, – в ряде случаев считают достаточным [19]. Потери при распространении радиоволн (РРВ) в диапазоне ММВ значительно превосходят потери в диапазоне ДМВ, поэтому оценку SINR для направленных радиолиний с учетом помех от сот только первого круга по модели на рисунке 1 можно считать обоснованной. Для имитационного моделирования помех в центральной соте СРД 5G, согласно ITU-R M.2135-1 [20] и ITU-R M.2412-0 [21], можно использовать модель из 7-ми секторизованных сот.

При ИМ направленных радиолиний для сценария диаграммообразования на основе позиционирования знание местоположения UE позволяет, *во-первых, настроить ориентацию луча* сектора базовой станции gNB в направлении на UE и, *во-вторых, настроить ширину луча* на данное UE так, чтобы минимизировать помехи для соседних устройств СРД. Совместная реализация этих процедур принципиально важна для повышения эффективности множественного доступа с SDMA.

Задачу ИМ совокупности направленных радиолиний в модели диаграммообразования на основе позиционирования можно сформулировать следующим образом. *Во-первых, требуется установить влияние ориентации и ширины луча сектора базовой станции gNB в направлении на UE на SINR. Во-вторых, необходимо оценить влияние погрешности определения местоположения UE при диаграммообразовании на основе позиционирования на SINR. В-третьих, следует выявить вклад в совокупное отношение SINR от помех, создаваемых радиолиниями SNOI по отдельности: а) внутри своего сектора, б) других секторов своей соты и в) других сот сети.* Установление данных зависимостей позволит количественно и качественно оценить возможности SDMA в сверхплотных СРД с направленными радиолиниями. Совокупность факторов погрешности позиционирования UE и ширины луча сектора базовой станции gNB при его ориентации в направлении на UE исследуются в комплексе. Гипотезой является утверждение о том, что чем точнее известны координаты UE, тем уже по азимуту и углу места может быть луч, сформированный сектором базовой станции gNB при диаграммообразовании на основе позиционирования для данного UE, при этом такой луч в условиях функционирования сверхплотной СРД может не перекрываться с другими лучами на соседние UE.

Для количественного и качественного исследования влияния ориентации и ширины луча сектора базовой станции gNB, а также погрешности определения местоположения (ОМП) пользовательского устройства на SDMA по критерию SINR, далее представлена функциональная схема имитационной модели совокупности радиолиний, работающих по принципу диаграммообразования LAB.

2.2. Функциональная схема модели совокупности радиолиний с LAB

Рисунок 2 иллюстрирует функциональную схему имитационной модели совокупности радиолиний с диаграммообразованием LAB.

В имитационной модели характеристика местоположения UE, включая текущую оценку координат и их погрешность (неопределенность), посту-

пает из *модуля позиционирования* UE сначала в модуль управления ориентацией луча по местоположению UE и затем в модуль управления шириной луча по местоположению UE. Местоположение пользовательского устройства характеризуется оценкой координат $\hat{\mathbf{x}}$.

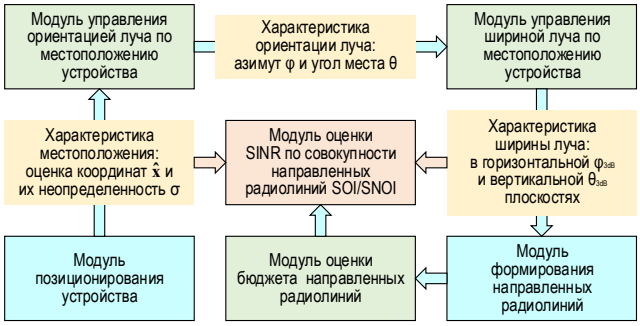


Рис. 2. Функциональная схема модели диаграммообразования LAB

Fig. 2. Location-Aware Beamforming Model Functional Diagram

По оценке координат $\hat{\mathbf{x}}$ пользовательского устройства UE_k обслуживаемого сектором s_i базовой станции gNB_j *модуль управления ориентацией луча* вычисляет необходимые направления ориентации диаграммы направленности антенны (ДНА) по азимуту $\varphi_{(\mathrm{i,k})}$ и углу места $\theta_{(i,k)}$ для каждого пользовательского устройства UE_k в секторе s_i модели сверхплотной СРД.

По неопределенности оценки координат σ пользовательского устройства UE_k *модуль управления шириной луча* вычисляет необходимую ширину луча по азимуту $\varphi_{3\mathrm{dB}(i,k)}$ и углу места $\theta_{3\mathrm{dB}(i,k)}$ для UE_k в секторе s_iмодели сверхплотной СРД. Ширина луча или ширина ДНА по уровню половинной мощности (–3 дБ) иногда обозначается параметром HPBW (*аббр. от англ.* Half-Power Beamwidth) и определяется в горизонтальной $\varphi_{3\mathrm{dB}(i,k)} = \mathrm{HPBW}_{H(i,k)}$ (Horizontal) и вертикальной $\theta_{3\mathrm{dB}(i,k)} = \mathrm{HPBW}_{\mathrm{V}(i,k)}$ (Vertical) плоскостях.

Модуль формирования направленных радиолиний выполняет их инициализацию в заданных направлениях ориентации по азимуту и углу места с заданной шириной луча в горизонтальной $\varphi_{3\mathrm{dB}(i,k)}$ и вертикальной $\theta_{3\mathrm{dB}(i,k)}$ плоскостях для каждого сектора s_i базовой станции gNB_j и пользовательского устройства UE_k в модели сверхплотной СРД.

В *модуле оценки бюджета направленных радиолиний* из набора $L_{(i,i,k)}$ выполняется расчет уровня принимаемого сигнала $P_{RX(j,i,k)}$ в каждой направленной радиолинии между сектором s_i базовой станции gNB_j и пользовательским устройством UE_k с учетом потерь PPB $\mathrm{PL}_{(i,i,k)}$ (*от англ* PathLoss) в модели сверхплотной СРД.

В модуле оценки SINR по совокупности направленных радиолиний SOI/SNOI выполняется оценка отношений $SINR_{(1,i,k)}$ по направленным радиолиниям из набора $L_{(1,i,k)}$ для 3-х секторов центральной соты базовой станции gNB_1 согласно сценарию на рисунке 1 для модели сверхплотной СРД. Бюджет направленных радиолиний SOI оценивается для набора $L_{(1,i,k)}$. Влияние помех SNOI на $\mathrm{SINR}_{(1,i,k)}$ оценивается по вкладу от радиолиний из набора: 1) $L_{(1,i,k\prime)}$, $k' \neq k$ внутри своего сектора s_i центральной соты базовой станции gNB_1; 2) $L_{(1,i\prime,k)}$ других секторов $s_{i\prime}$, $i' \neq i$ своей соты базовой станции gNB_1; 3) $L_{(j,i,k)}$ 3-х секторов s_i окружающих сот gNB_j.

Формализуем далее функции каждого модуля имитационной модели совокупности радиолиний с диаграммообразованием по схеме на рисунке 2.

2.3. Модули формирования и обработки совокупности радиолиний с LAB

2.3.1. Модуль позиционирования устройства

В спецификации 3GPP TS 23.273 [22] для сетей 5G NR (*аббр. от англ.* New Radio) формализована организация сервисов позиционирования (LCS, *аббр. от англ.* Location Services). Позиционирование UE радиотехническими методами основано, преимущественно, на разностно-дальномерных и угломерных первичных измерениий и использует метод оценки разности времен прихода сигнала в канале «вниз» (DL-TDOA, *аббр. от англ.* Downlink Time Difference of Arrival) и метод оценки угла прихода сигнала в канале «вверх» (UL-AOA, *аббр. от англ.* Angle of Arrival).

Рисунок 3 иллюстрирует характеристику местоположения устройства на плоскости, где $\mathbf{x} = [x, y, z]$ – координаты истинного местоположения UE; $\hat{\mathbf{x}} = [\hat{x}, \hat{y}, \hat{z}]$ – оценка координат (ОК) местоположения UE; σ – неопределенность оценки местоположения, задающая окружность диаметра σ с центром в точке $\hat{\mathbf{x}} = [\hat{x}, \hat{y}, \hat{z}]$.

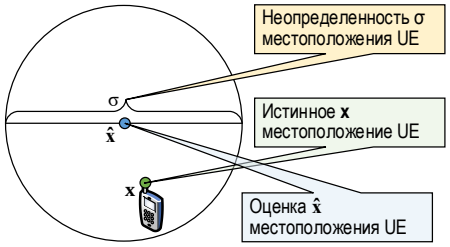


Рис. 3. Характеристика местоположения UE

Fig. 3. UE Location Characteristic

Согласно спецификации 3GPP TS 22.071 [23] каждому LCS соответствует своя категория горизонтальной точности, которая может быть представлена одной из фигур географической протяженности (GAD, *аббр. от англ.* Geographical Area Description), формализованной в 3GPP TS 23.032 [24]. Неопределенность местоположения характеризует погрешности ОК кругом диаметром σ с центром в точке $\hat{\mathbf{x}} = [\hat{x}, \hat{y}, \hat{z}]$. Функциональным результатом работы модуля позиционирования устройства является характеристика местоположения UE, вклю-

чающая оценку координат $\hat{\mathbf{x}} = [\hat{x}, \hat{y}, \hat{z}]$ и их неопределенность σ.

В 3GPP TS 22.261 [25] специфицированы категории неопределенности местоположения, обозначенные индексами PSL (*аббр. от англ.* Positioning Service Level). Каждому из 6 индексов PSL соответствует неопределенность абсолютного местоположения σ UE в горизонтальной плоскости вместе с другими ключевыми показателями эффективности KPI (*аббр. от англ.* Key Performance Indicator) и сценариями функционирования СРД 5G. Таблица 1 содержит параметр неопределенности σ для сценариев сетей 5G согласно 3GPP TS 22.261 [25].

ТАБЛИЦА 1. Параметр неопределенности местоположения

TABLE 1. Location Uncertainty Parameter

Индекс PSL	Покрытие	Сценарий	σ/2, м
1	внутри/снаружи помещений	село/город	10
2	снаружи помещений	село/город/ плотный город	3
3,4	снаружи помещений	село/город/ плотный город	1
5,6	снаружи помещений	село/ плотный город	0,3

Исследование влияния точности позиционирования на отношение SINR будем выполнять в диапазоне σ от 10 до 1 м. Далее формализуем функции модуля управления ориентацией луча по местоположению устройства.

2.3.2. Модуль управления ориентацией луча по местоположению устройства

Рассмотрим на рисунке 4 зону обслуживания сектора s_i трехсекторной базовой станции gNB_j и точку $\hat{\mathbf{x}}_k$ ОК пользовательского устройства UE_k. Согласно сделанному ранее допущению, АР каждого сектора s_i базовой станции позволяет одновременно формировать несколько лучей при работе на передачу и обслуживать несколько пользовательских устройств UE_k. Ориентация каждого луча АР сектора s_i выполняется в направлении на каждое пользовательское устройство UE_k по известной оценке его координат $\hat{\mathbf{x}}_k$. Будем полагать (см. рисунок 4), что точка $\hat{\mathbf{x}}_k$ в зоне обслуживания сектора s_i характеризуется углом азимута $\varphi_{(i,k)}$ в горизонтальной плоскости и углом места $\theta_{(i,k)}$ в вертикальной плоскости. Углы азимута $\varphi_{(i,k)}$ измеряются относительно центра сектора s_i в диапазоне от 0 ° до 360 °, а углы места $\theta_{(i,k)}$ измеряются относительно центра сектора s_i в диапазоне от –90 ° до 90 °.

Характеристика ориентации луча сектора s_i на каждое пользовательское устройство UE_k включает направление (*steering*) по азимуту $\varphi_{(i,k)}$ в горизонтальной плоскости (см. рисунок 4a) и наклон (*tilting*) по углу места $\theta_{(i,k)}$ в вертикальной плоскости (см. рисунок 4b).

Далее введем обозначения для характеристики ширины луча на рисунке 4. Будем полагать, что *ширина луча моделируется конусом и измеряется по уровню –3 дБ от максимума* в заданной ориентации по азимуту $\varphi_{(i,k)}$ и углу места $\theta_{(i,k)}$. В горизонтальной плоскости обозначим ширину луча через $\varphi_{3\text{dB}(i,k)}$, а в вертикальной – через $\theta_{3\text{dB}(i,k)}$.

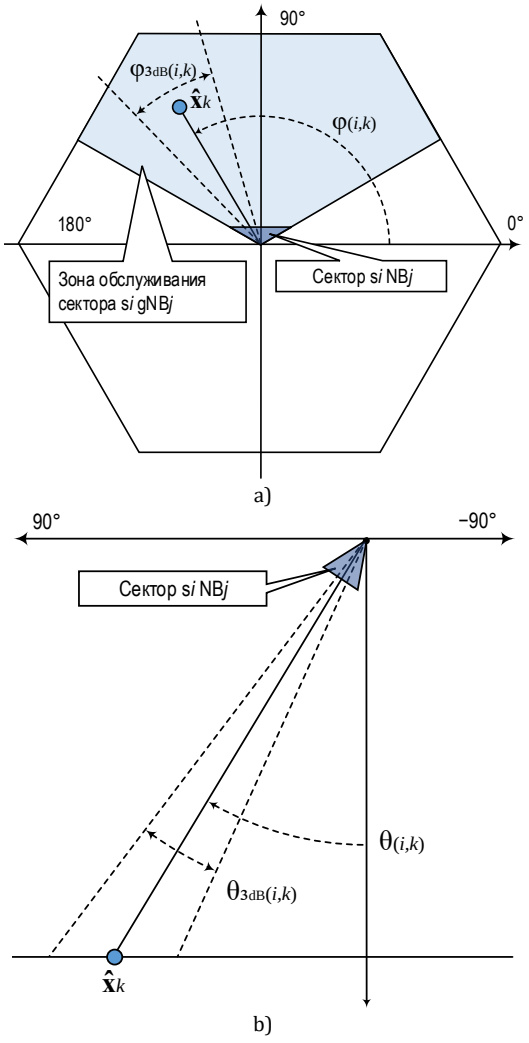


Рис. 4. Характеристика ориентации луча на пользовательское устройство: a) по азимуту (ширина луча в горизонтальной плоскости); b) по углу места (ширина луча в вертикальной плоскости)

Fig. 4. Characteristics of Beam Orientation to the UE: a) in Azimuth (Beamwidth in the Horizontal Plane); b) by Elevation (Beamwidth in the Vertical Plane)

Характеристика ширины луча сектора s_i на каждое пользовательское устройство UE_k включает ширину луча $\varphi_{3\text{dB}(i,k)}$ в горизонтальной плоскости (см. рисунок 4a) и ширину луча $\theta_{3\text{dB}(i,k)}$ в вертикальной (см. рисунок 4b).

Из рисунка 4 следует, что ширина луча в горизонтальной $\varphi_{3\text{dB}(i,k)}$ и вертикальной $\theta_{3\text{dB}(i,k)}$ плоскостях существенно влияет на зону покрытия АР сектора s_i данным лучом, что, в свою очередь, определяет уровень принятого сигнала не только в точке $\hat{\mathbf{x}}_k$, на которую ориентирован данный луч, но и в

окрестности данной точки. Зона покрытия лучом, который сориентирован в точку $\hat{\mathbf{x}}_k$, количественно характеризуется кортежем из четырех углов $\varphi_{(i,k)}, \theta_{(i,k)}, \varphi_{3\text{dB}(i,k)}, \theta_{3\text{dB}(i,k)}$. Если для настройки ориентации луча достаточно оценки координат $\hat{\mathbf{x}}_k$, то для настройки ширины луча, необходимой для радиопокрытия заданной точки $\hat{\mathbf{x}}_k$, АР сектора s_i необходимо знать неопределенность σ местоположения $\hat{\mathbf{x}}_k$. Далее формализуем функции модуля управления шириной луча по неопределенности местоположения устройства.

2.3.2. Модуль управления шириной луча по местоположению устройства

Рисунок 5 иллюстрирует порядок настройки ширины луча в горизонтальной $\varphi_{3\text{dB}(i,k)}$ и вертикальной $\theta_{3\text{dB}(i,k)}$ плоскостях по оценке местоположения $\hat{\mathbf{x}}_k$ пользовательского устройства UE_k и его неопределенности σ для АР сектора s_i [14]. Показана зона обслуживания одного сектора s_i базовой станции gNB_j и оценка координат $\hat{\mathbf{x}}_k$ пользовательского устройства UE_k, неопределенность местоположения которой характеризуется параметром σ.

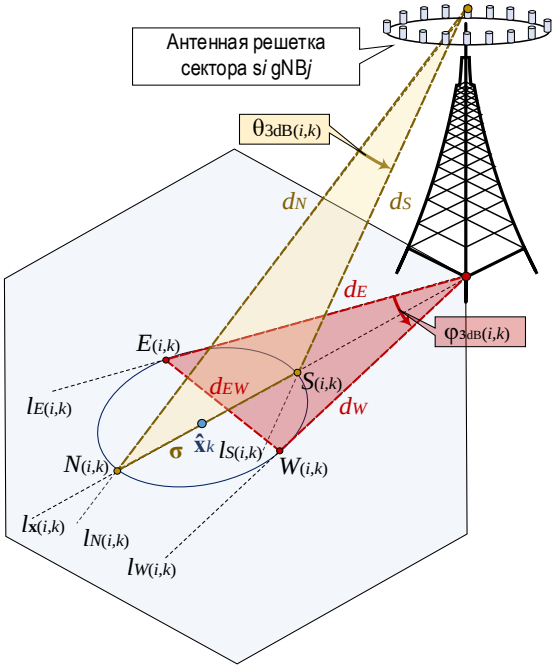


Рис. 5. Настройка ширины луча в горизонтальной и вертикальной плоскостях

Fig. 5. Beamwidth Tuning in the Horizontal and Vertical Planes

Будем полагать, что ориентация АР сектора s_i в глобальной системе координат определяется относительно центра сектора s_i следующим образом (см. рисунок 5): север (North) – сверху; юг (South) – снизу; запад (West) – слева; восток (East) – справа.

Рассмотрим два луча $l_{E(i,k)}$ и $l_{W(i,k)}$, исходящие из точки $\boldsymbol{s}_i^H$ проекции центра сектора s_i в горизонтальной плоскости и касательные точки на окружности неопределенности местоположения UE_k диаметра σ с центром в точке $\hat{\mathbf{x}}_k$. Обозначим касательную точку луча $l_{E(i,k)}$ и окружности неопределенности местоположения UE_k через $E_{(i,k)}$ (восток), а касательную точку луча $l_{W(i,k)}$ и окружности неопределенности местоположения UE_k через $W_{(i,k)}$ (запад). Угол между двумя лучами l_E и l_W в горизонтальной плоскости представляет собой ширину АР сектора s_i по азимуту $\varphi_{3\text{dB}(i,k)}$ для пользовательского устройства UE_k.

Рассмотрим теперь луч $l_{\mathbf{x}(i,k)}$, исходящий из точки $\boldsymbol{s}_i^H$ проекции центра сектора s_i в горизонтальной плоскости, и проходящий через точку $\hat{\mathbf{x}}_k$ оценки местоположения UE_k. Данный луч $l_{\mathbf{x}(i,k)}$ пересекает окружность неопределенности местоположения UE_k в точках $N_{(i,k)}$ (север) и $S_{(i,k)}$ (юг). Проведем два луча $l_{N(i,k)}$ и $l_{S(i,k)}$, исходящие из точки $\boldsymbol{s}_i^V$ сектора s_i в вертикальной плоскости, и касательные к окружности неопределенности местоположения UE_k диаметра σ с центром в точке $\hat{\mathbf{x}}_k$ в вертикальной плоскости. Обозначим касательную точку луча $l_{N(i,k)}$ и окружности неопределенности местоположения UE_k через $N_{(i,k)}$ (север), а касательную точку луча $l_{S(i,k)}$ и окружности неопределенности местоположения UE_k через $S_{(i,k)}$ (юг). Угол между двумя лучами l_N и l_S в вертикальной плоскости представляет собой ширину луча АР сектора s_i по углу места $\theta_{3\text{dB}(i,k)}$ для пользовательского устройства UE_k.

Таким образом, из приведенного на рисунке 5 порядка настройки ширины луча можно сделать *качественный* вывод о том, что необходимая и достаточная для радиопокрытия пользовательского устройства UE_k ширина луча по уровню половинной мощности HPBW непосредственно определяется погрешностью σ оценки его координат $\hat{\mathbf{x}}_k$. *Количественную* оценку требуемой ширины луча можно выполнить по теореме косинусов из анализа двух треугольников.

Рассмотрим на рисунке 5 треугольник в горизонтальной плоскости, образованный сторонами d_E, d_W и d_{EW}. *Ширину луча в горизонтальной плоскости* $\varphi_{3\text{dB}(i,k)}$ для направленной радиолинии (i,k) между АР сектора s_i и пользовательским устройством UE_k с учетом равенства сторон $d_E = d_W$ можно определить выражением:

$$\varphi_{3\text{dB}(i,k)} = \arccos\left(1 - \frac{d_{EW}^2}{2d_E^2}\right). \qquad (1)$$

Рассмотрим треугольник в вертикальной плоскости, образованный сторонами d_N, d_S и σ (см рисунок 5). *Ширину луча в вертикальной плоскости* $\theta_{3\text{dB}(i,k)}$ для направленной радиолинии (i,k) между АР сектора s_i и пользовательским устройством UE_k можно определить выражением:

$$\theta_{3\text{dB}(i,k)} = \arccos\left(\frac{d_N^2 + d_S^2 - \sigma^2}{2d_N d_S}\right). \qquad (2)$$

Из анализа порядка настройки ширины луча в горизонтальной и вертикальной плоскостях (см. рисунок 5) следует, что ориентация и ширина диаграммы направленности АР адаптируется к текущему местоположению $\hat{\mathbf{x}}_k$ и его неопределенности σ. С точки зрения влияния удаленности UE и неопределенности его местоположения характер адаптации можно охарактеризовать следующим образом: 1) чем ближе пользовательское устройство UE_k располагается к обслуживающему сектору s_i, тем шире получится луч; и наоборот, чем дальше UE_k располагается от s_i, например, на границе обслуживающего сектора, тем у́же получится луч; 2) чем меньше неопределенность σ оценки координат $\hat{\mathbf{x}}_k$, тем у́же в горизонтальной и вертикальной плоскостях будет луч; и наоборот, чем больше неопределенность σ оценки координат $\hat{\mathbf{x}}_k$, тем шире в горизонтальной и вертикальной плоскостях будет луч. Далее формализуем алгоритм работы модуля управления шириной луча по местоположению для совокупности радиолиний.

2.3.4. Алгоритм управления шириной луча по местоположению устройства

Управление шириной луча по местоположению осуществляется для совокупности направленных радиолиний (i,k) в каждом секторе s_i базовой станции $\mathrm{gNB}_j, j = 1,\ldots,7$ модели сверхплотной СРД (см. рисунок 1). Каждая направленная радиолиния (i,k) количественно характеризуется кортежем из четырех углов $\varphi_{(i,k)}$, $\theta_{(i,k)}$, $\varphi_{3\mathrm{dB}(i,k)}$, $\theta_{3\mathrm{dB}(i,k)}$. Технологически минимальная ширина ДНА по уровню половинной мощности HPBW определяется конструктивом и размерностью АР, поэтому для имитационной модели необходимо инициализировать минимальную ширину луча в горизонтальной $\varphi_{3\mathrm{dBmin}}$ и вертикальной $\theta_{3\mathrm{dBmin}}$ плоскостях. Так, для эквидистантной прямоугольной АР (URA, *аббр. от англ.* Uniform Rectangular Array) известна следующая оценка ширины луча диаграммы направленности по уровню половинной мощности в горизонтальной и вертикальной плоскостях [26–29]:

$$\mathrm{HPBW}_H = \mathrm{HPBW}_V \approx \frac{1{,}772}{N-1} \text{ рад}, \qquad (3)$$

где N – число элементов АР вдоль оси $0x$ и вдоль оси $0y$ при расстоянии между элементами в половину длины волны. Например, при размерности 32×32 ширина луча ДНА будет составлять $\varphi_{3\mathrm{dB}} = \theta_{3\mathrm{dB}} = 3{,}3\,°$.

Скрипт 1 содержит процедуры алгоритма работы модуля диаграммообразования на основе позиционирования [14].

Входными данными алгоритма являются: набор $i \in \mathbb{I}$ секторов s_i; набор $k \in \mathbb{K}$ оценок координат $\hat{\mathbf{x}}_k$ пользовательских устройств UE_k; неопределенность σ оценок координат; минимальная ширина луча в горизонтальной и вертикальной плоскостях $\varphi_{3\mathrm{dBmin}} = \theta_{3\mathrm{dBmin}} = 3\,°$.

Скрипт 1. Алгоритм модуля диаграммообразования LAB

1 Входные параметры: $\mathbb{J}, \mathbb{I}, \mathbb{K}, \sigma, \varphi_{3\mathrm{dBmin}}, \theta_{3\mathrm{dBmin}}$
2 Выходные параметры: $L_{(j,i,k)}, \varphi_{(i,k)}, \theta_{(i,k)}, \varphi_{3\mathrm{dB}(i,k)}, \theta_{3\mathrm{dB}(i,k)}$
3 Цикл по сотам j в наборе $\mathbb{J}$
4 Цикл по секторам i в наборе $\mathbb{I}$
5 Цикл по устройствам k в наборе $\mathbb{K}$
6 // Инициализация направленных радиолиний набора $L_{(j,i,k)}$
7 Оценка расстояний $\mathrm{gNB}_{js_i} \to \mathrm{UE}_k$ в 2D
8 // Вычисление касательных к окружности неопределенности
9 $[W_{(i,k)}, E_{(i,k)}, N_{(i,k)}, S_{(i,k)}] = f\begin{pmatrix} \boldsymbol{s}_i^H, \boldsymbol{s}_i^V, \hat{\mathbf{x}}_k, \sigma, l_{\mathbf{x}(i,k)}, \ldots \\ l_{E(i,k)}, l_{W(i,k)}, l_{N(i,k)}, l_{S(i,k)} \end{pmatrix}$
10 // Настройка ширины луча в горизонтальной плоскости
11 $d_E = \boldsymbol{s}_i^H \overset{H}{\to} E_{(i,k)}$; $d_W = \boldsymbol{s}_i^H \overset{H}{\to} W_{(i,k)}$; $d_E = d_W$
12 $d_{EW} = W_{(i,k)} \overset{H}{\to} E_{(i,k)}$
13 $\varphi_{3\mathrm{dB}(i,k)} = \arccos\left(1 - \frac{d_{EW}^2}{2d_E^2}\right)$
14 $\varphi_{3\mathrm{dB}(i,k)} = \max\left(\varphi_{3\mathrm{dB}(i,k)}, \varphi_{3\mathrm{dBmin}}\right)$
15 // Настройка ширины луча в вертикальной плоскости
16 $d_N = \boldsymbol{s}_i^V \overset{V}{\to} N_{(i,k)}$
17 $d_S = \boldsymbol{s}_i^V \overset{V}{\to} S_{(i,k)}$
18 $\theta_{3\mathrm{dB}(i,k)} = \arccos\left(\frac{d_N^2 + d_S^2 - \sigma^2}{2d_N d_S}\right)$
19 $\theta_{3\mathrm{dB}(i,k)} = \max\left(\theta_{3\mathrm{dB}(i,k)}, \theta_{3\mathrm{dBmin}}\right)$
20 Завершение цикла по устройствам k в наборе $\mathbb{K}$
21 Завершение цикла по секторам i в наборе $\mathbb{I}$
22 Завершение цикла по сотам j в наборе $\mathbb{J}$

Выходными данными алгоритма является набор $L_{(j,i,k)}$ совокупности направленных радиолиний модели сверхплотной СРД между i секторами из набора $i \in \mathbb{I}$ и k пользовательскими устройствами UE_k с оценками координат $\hat{\mathbf{x}}_k$ из набора $k \in \mathbb{K}$ для каждой базовой станции gNB_j из набора $j \in \mathbb{J}$. Для каждой радиолинии $\mathrm{gNB}_{js_i} \to \mathrm{UE}_k$ из набора $L_{(j,i,k)}$ совокупности направленных радиолиний записывается кортеж из четырех углов $\varphi_{(i,k)}$, $\theta_{(i,k)}$, $\varphi_{3\mathrm{dB}(i,k)}$, $\theta_{3\mathrm{dB}(i,k)}$.

Строки 3–5 начинают, а строки 20–22 завершают цикл алгоритма (скрипт 1) по сотам j в наборе $\mathbb{J}$, секторам i в наборе $\mathbb{I}$ и точкам ОК устройств k в наборе $\mathbb{K}$, соответственно. При инициализации каждой направленной радиолинии (i,k) для UE_k в секторе s_i соты базовой станции gNB_j выполняется оценка расстояний между точкой проекции центра сектора $\boldsymbol{s}_i^H$ в горизонтальной плоскости и точкой ОК $\hat{\mathbf{x}}_k$ устройства UE_k:

$$d_{2D(i,k)} = \|\boldsymbol{s}_i^H - \hat{\mathbf{x}}_k\|, \qquad (4)$$

где $\boldsymbol{s}_i^H = [x_i, y_i, z_i]$ – координаты точки проекции центра сектора s_i в горизонтальной плоскости, т. е. $z_i = 0$; $\hat{\mathbf{x}}_k = [\hat{x}_k, \hat{y}_k, \hat{z}_k]$ – оценка координат пользовательского устройства UE_k; $\|\cdot\|$ – оператор нормы вектора в евклидовом пространстве, определяемый выражением:

$$\|\boldsymbol{s}_i^H - \hat{\mathbf{x}}_k\| = \\ = \sqrt{(x_i - \hat{x}_k)^2 + (y_i - \hat{y}_k)^2 + +(z_i - \hat{z}_k)^2}. \qquad (5)$$

Далее выполняется вычисление касательных точек $W_{(i,k)}, E_{(i,k)}, N_{(i,k)}, S_{(i,k)}$, полученных в результате касания окружности неопределенности местоположения лучами $l_{E(i,k)}, l_{W(i,k)}, l_{N(i,k)}, l_{S(i,k)}$, соответственно. Оценка значений $\varphi_{3\text{dB}(i,k)}$ и $\theta_{3\text{dB}(i,k)}$ выполняется при ограничении на минимальную ширину луча $\varphi_{3\text{dBmin}}, \theta_{3\text{dBmin}}$.

Последовательность оценки $\varphi_{3\text{dB}(i,k)}$ включает построение треугольника из отрезков d_W, d_E и d_{EW}. Отрезок d_E получается в результате соединения точки $\boldsymbol{s}_i^H$ с точкой $E_{(i,k)}$ лучом $l_{E(i,k)}$: $\boldsymbol{s}_i^H \overset{H}{\rightarrow} E_{(i,k)}$; оператор $(\cdot) \overset{H}{\rightarrow} (\cdot)$ обозначает соединение ребра через две вершины в горизонтальной плоскости.

Аналогично отрезок d_W получается в результате соединения точки $\boldsymbol{s}_i^H$ с точкой $W_{(i,k)}$ лучом $l_{W(i,k)}$: $\boldsymbol{s}_i^H \overset{H}{\rightarrow} W_{(i,k)}$.

Отрезок d_{EW} получается в результате соединения точки $W_{(i,k)}$ с точкой $E_{(i,k)}$: $d_{EW} = W_{(i,k)} \overset{H}{\rightarrow} E_{(i,k)}$. Далее по полученным сторонам треугольника d_W, d_E и d_{EW} выполняется оценка $\varphi_{3\text{dB}(i,k),}$ согласно (1).

Последовательность оценки $\theta_{3\text{dB}(i,k)}$ включает построение треугольника из отрезков d_N, d_S и σ:

– отрезок d_N получается в результате соединения точки $\boldsymbol{s}_i^V$ с точкой $N_{(i,k)}$ лучом $l_{N(i,k)}$: $\boldsymbol{s}_i^V \overset{V}{\rightarrow} N_{(i,k)}$;

– отрезок d_S получается в результате соединения точки $\boldsymbol{s}_i^V$ с точкой $S_{(i,k)}$ лучом $l_{S(i,k)}$: $\boldsymbol{s}_i^V \overset{V}{\rightarrow} S_{(i,k)}$;

Далее по полученным сторонам треугольника d_N, d_S и σ выполняется оценка $\theta_{3\text{dB}(i,k)}$, согласно (2). После чего формализуем функции модуля формирования направленных радиолиний в модели сверхплотной СРД.

2.3.5. Модуль формирования направленных радиолиний с LAB

Диаграммообразование LAB осуществляется для совокупности направленных радиолиний (i,k) в каждом секторе s_i базовой станции gNB_j модели сверхплотной СРД (см. рисунок 1).

Каждая направленная радиолиния (i,k) сектора s_i количественно характеризуется кортежем из четырех углов $\varphi_{(i,k)}$, $\theta_{(i,k)}$, $\varphi_{3\text{dB}(i,k)}$, $\theta_{3\text{dB}(i,k)}$. Далее назовем такой кортеж трафиковым лучом (i,k). Набор трафиковых лучей $L_{(i,k)}$ представляет собой исходные данные для модуля оценки бюджета направленных радиолиний и последующей их классификации на направленные радиолинии SOI и направленные радиолинии помех SNOI.

Введем *ограничения и допущения работы модуля формирования направленных радиолиний при организации* узких трафиковых лучей в настоящей имитационной модели: 1) каждая базовая станция gNB_j обслуживает три неперекрывающихся сектора s_i; при этом каждый сектор s_i обслуживается своей АР; 2) в каждом секторе s_i обслуживаются K пользовательских устройств UE_k посредством индивидуальных трафиковых лучей из набора $L_{(i,k)}$; 3) на секторах s_i базовых станций gNB_j предполагается обеспечение радиопокрытия в каждом секторе s_i соты gNB_j; 4) каждый сектор s_i соты gNB_j осведомлен об оценке $\hat{\mathbf{x}}_k$ текущего местоположения пользовательского устройства UE_k в своей зоне обслуживания; 5) АР каждого сектора s_i соты gNB_j образована из достаточно большого числа излучающих элементов N^2, которое значительно больше числа K пользовательских устройств UE_k, обслуживаемых в данном секторе s_i; 6) каждое пользовательское устройство UE_k может быть обслужено достаточно узким отдельным трафиковым лучом, формируемым АР сектора s_i; 7) каждый отдельный трафиковый луч задается кортежем $\varphi_{(i,k)}, \theta_{(i,k)}, \varphi_{3\text{dB}(i,k)}, \theta_{3\text{dB}(i,k)}$ и характеризуется ориентацией и шириной в заданном секторе s_i; 8) анализ показателей пространственного уплотнения одновременных передач SDMA для совокупности направленных радиолиний по критерию SINR производится для набора одновременно работающих трафиковых лучей $L_{(i,k)}$ каждого сектора s_i; 9) трафиковые лучи на каждое пользовательское устройство UE_k из набора $L_{(j,i,k)}$, всех сот сети активируются одновременно во всех секторах s_i всех базовых станций gNB_j; 10) в каждом узком луче радиолинии $\text{gNB}_{js_i} \rightarrow \text{UE}_k$ передача ведется с максимальной мощностью, т. е. адаптация мощности в трафиковом канале «вниз» не используется; 11) ориентация луча АР каждого сектора s_i осуществляется на известную сети точку оценки координат $\hat{\mathbf{x}}_k$ пользовательского устройства UE_k.

Далее формализуем функции модуля оценки бюджета потерь в направленных радиолиниях модели сверхплотной СРД.

2.3.6. Модуль оценки бюджета направленных радиолиний с LAB

Оценим уровень сигнала, принятого в точке истинного (***true***) местоположения $\mathbf{x}_k$ пользовательского устройства UE_k, при ориентации луча на точку оценки (***estimate***) местоположения $\hat{\mathbf{x}}_k$ пользовательского устройства UE_k. Далее для краткости обозначим сектор s_i символом i, индекс точки $\mathbf{x}_k$ символом k_t (***true***), а индекс точки $\hat{\mathbf{x}}_k$ символом k_e (***estimate***). Будем полагать, что UE_k располагается в дальней зоне излучения АР сектора s_i. Тогда мощность сигнала $P^{RX}_{(i,k_e,k_t)}$, принятого UE_k в точке $\mathbf{x}_k$ истинного местоположения с индексом k_t от луча сектора s_i, ориентированного в точку $\hat{\mathbf{x}}_k$ оценки местоположения с индексом k_e можно оценить по формуле [17]:

$$P^{RX}_{(i,k_e,k_t)} = P^{TX}_S - L^{PL}_{(i,k_t)} + \underbrace{+A^{AZ}_{(i,k_e,k_t)} + A^{EL}_{(i,k_e,k_t)}}_{\text{ДНА}} + G^{TX}_S + \underbrace{+B^{AZ}_{(i,k_e,k_t)} + B^{EL}_{(i,k_e,k_t)} + G^{BF}_S}_{\text{КУ при ДО}}, \quad (6)$$

где P^{TX}_S – максимальная мощность передачи всей АР, расположенной в центре сектора s_i; $L^{PL}_{(i,k_t)}$ – потери при РРВ в пространстве между точкой s_i и точкой $\mathbf{x}_k$ с индексом k_t; $A^{AZ}_{(i,k_e,k_t)}$ и $A^{EL}_{(i,k_e,k_t)}$ – ДНА антенной решетки сектора s_i в горизонтальной и вертикальной плоскостях, определенные в (10) и (11), соответственно; G^{TX}_S – максимальный коэффициент усиления (КУ) одного элемента АР, расположенной в центре сектора s_i; G^{BF}_S – максимальный КУ при диаграммообразовании (ДО); $B^{AZ}_{(i,k_e,k_t)}$ и $B^{EL}_{(i,k_e,k_t)}$ – оценка КУ АР в точке $\mathbf{x}_k$ с индексом k_t при диаграммообразовании на точку $\hat{\mathbf{x}}_k$ с индексом k_e [16]:

$$B^{AZ}_{(i,k_e,k_t)} = 10\log_{10}\left[\mathrm{sinc}\left(\frac{\varphi_{(i,k_t)} - \varphi_{(i,k_e)}}{1{,}13\cdot\varphi_{3\mathrm{dB}(i,k_e)}}\right)^2\right], \quad (7)$$

$$B^{EL}_{(i,k_e,k_t)} = 10\log_{10}\left[\mathrm{sinc}\left(\frac{\theta_{(i,k_t)} - \theta_{(i,k_e)}}{1{,}13\cdot\theta_{3\mathrm{dB}(i,k_e)}}\right)^2\right]. \quad (8)$$

Из анализа (6–8) следует, что мощность принятого сигнала вычисляется из набора параметров, которые масштабируют мощность передачи АР.

Нормированный коэффициент ослабления сигнала в точке k_t по сравнению с точкой k_e, на которую ориентирован луч, можно выразить формулой:

$$F_{(i,k_e,k_t)} = \left(10^{\frac{A^{AZ}_{(i,k_e,k_t)} + A^{EL}_{(i,k_e,k_t)}}{10}}\right)^2, \quad (9)$$

где $A^{AZ}_{(i,k_e,k_t)}$ и $A^{EL}_{(i,k_e,k_t)}$ – диаграмма направленности АР в дБ в горизонтальной (по азимуту AZ) и вертикальной (по углу места EL) плоскостях, наблюдаемая в точке $\mathbf{x}_k$ с индексом k_t при ориентации луча от центра сектора s_i в точку $\hat{\mathbf{x}}_k$ с индексом k_e. ДНА в горизонтальной и вертикальной плоскостях, наблюдаемые по азимуту $\varphi_{(i,k_t)}$ и углу места $\theta_{(i,k_t)}$ в точке $\mathbf{x}_k$ с индексом k_t при ориентации луча по азимуту $\varphi_{(i,k_e)}$ и углу места $\theta_{(i,k_e)}$ в точку $\hat{\mathbf{x}}_k$ с индексом k_e, можно определить выражениями [20, 21]:

$$A^{AZ}_{(i,k_e,k_t)} = -\min\left[12\left(\frac{\varphi_{(i,k_t)} - \varphi_{(i,k_e)}}{\varphi_{3\mathrm{dB}(i,k_e)}}\right), A^{AZ}_{\min}\right], \quad (10)$$

$$A^{EL}_{(i,k_e,k_t)} = -\min\left[12\left(\frac{\theta_{(i,k_t)} - \theta_{(i,k_e)}}{\theta_{3\mathrm{dB}(i,k_e)}}\right), A^{EL}_{\min}\right], \quad (11)$$

где $A^{AZ}_{\min}$ и $A^{EL}_{\min}$ – ограничение на уровень боковых лепестков по азимуту и углу места, соответственно. Из анализа выражений (10) и (11) следует, что максимум ДНА получится при $\varphi_{(i,k_t)} = \varphi_{(i,k_e)}$ и $\theta_{(i,k_t)} = \theta_{(i,k_e)}$, т. е. тогда, когда точка $\hat{\mathbf{x}}_k$ с индексом k_t ОК UE_k, используемая при настройке ориентации и ширины луча, совпадает с точкой $\mathbf{x}_k$ с индексом k_t истинного местоположения UE_k. При этом значения ширины луча по азимуту $\varphi_{3\mathrm{dB}(i,k_e)}$ и углу места $\theta_{3\mathrm{dB}(i,k_e)}$ выступают масштабирующими параметрами диаграммы направленности: чем шире луч, тем бóльшую территорию он покрывает и, следовательно, слабее оказывается влияние ориентации по азимуту и углу места.

Рисунок 6 иллюстрирует сценарий оценки бюджета потерь в направленных радиолиниях с прямой видимостью (LOS, *аббр. от англ.* Line-of-Sight) согласно 3GPP TR 38.901 [31] для микросоты улиц городского каньона (UMi, *аббр. от англ.* Urban Micro – Street Canyon). Оценка потерь в радиолиниях LOS выполняется по формуле [31]:

$$L^{PL}_{(i,k_t)} = 32{,}4 + 21\lg(d_{3D}) + 20\lg(f_c), \quad (12)$$

где d_{3D} – 3D расстояние между центром сектора s_i базовой станции gNB и местоположением UE_k с учетом высоты h_{gNB} подвеса АР сектора s_i и высоты h_{UE} антенны UE_k; d_{2D} – 2D расстояние между gNB и UE_k; f_c – несущая частота в Гц.

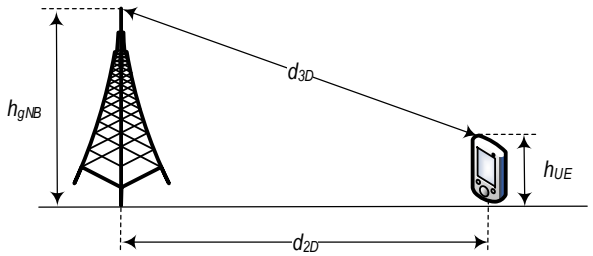


Рис. 6. Сценарий оценки бюджета потерь в направленных радиолиниях LOS

Fig. 6. Scenario for Directional Link Budget Estimate in LOS

Оценка расстояния между точкой центра сектора $\boldsymbol{s}^V_i$ в пространстве и точкой ОК $\hat{\mathbf{x}}_k$ устройства UE_k в пространстве можно оценить по формуле:

$$d_{3D(i,k)} = \left\|\boldsymbol{s}^V_i - \hat{\mathbf{x}}_k\right\|; \quad (13)$$

где $\boldsymbol{s}^V_i = [x_i, y_i, z_i]$ – координаты точки центра сектора s_i в пространстве.

Далее формализуем функции модуля оценки SINR по совокупности радиолиний SOI/SNOI.

2.3.7. Модуль оценки SINR по совокупности радиолиний SOI/SNOI

В имитационной модели анализируется набор из точек $k_e \in \mathbb{K}_e$ ОК $\hat{\mathbf{x}}_k$ пользовательского устройства UE_k и набор из точек $k_t \in \mathbb{K}_t$ истинных местоположений $\mathbf{x}_k$ пользовательского устройства UE_k. *Ориентация луча* каждого сектора s_i на каждое пользовательское устройство UE_k в имитационной модели осуществляется по точкам $\hat{\mathbf{x}}_k$ оценок координат UE_k из набора $k_e \in \mathbb{K}_e$.

Оценка уровня принимаемого SOI в имитационной модели осуществляется по точкам $k_t \in \mathbb{K}_t$ и

$k_e \in \mathbb{K}_e$ из набора $L_{(1,i,k)}$ в трех секторах центральной соты базовой станции gNB_1. Оценка уровня принимаемых помех SNOI в ИМ также осуществляется по точкам $k_t \in \mathbb{K}_t$ и $k_e \in \mathbb{K}_e$ из набора $L_{(1,i,k)}$ в трех секторах центральной соты базовой станции gNB_1. При этом в каждой точке истинных местоположений $\mathbf{x}_k$ и оценок координат $\hat{\mathbf{x}}_k$ из набора $L_{(1,i,k)}$ учитывается вклад помех от одновременной работы трафиковых лучей направленных радиолиний SNOI из набора $L_{(j,i,k)}$, от всех остальных сот и секторов рассматриваемой модели сверхплотной сети радиодоступа.

В результате моделирования получаем два набора значений $SINR_{(1,i,k_e,k_t)}$ и $SINR_{(1,i,k_e,k_e)}$.

Набор значений $SINR_{(1,i,k_e,k_t)}$ рассчитывается в трех секторах центральной соты базовой станции gNB_1 для точек $\mathbf{x}_k$ истинных местоположений пользовательских устройств UE_k из набора $k_t \in \mathbb{K}_t$ при ориентации трафиковых лучей на точки $\hat{\mathbf{x}}_k$ оценок координат UE_k из набора $k_e \in \mathbb{K}_e$.

Набор значений $SINR_{(1,i,k_e,k_e)}$ рассчитывается в трех секторах центральной соты базовой станции gNB_1 для точек $\hat{\mathbf{x}}_k$ оценок UE_k из набора $k_e \in \mathbb{K}_e$ при ориентации лучей на точки $\hat{\mathbf{x}}_k$ оценок координат UE_k из набора $k_e \in \mathbb{K}_e$.

При одновременной работе трафиковых лучей всех направленных радиолиний из набора $L_{(j,i,k)}$, рассматриваемой модели сверхплотной СРД отношение $SINR_{(1,i,k_e,k_t)}$ можно оценить в виде (14), где $P^{RX}_{(1,i,k_e,k_t)}$ – мощность принятого SOI (6) для трех секторов центральной соты базовой станции gNB_1; P_N – мощность шума.

$$SINR_{(1,i,k_e,k_t)} = \frac{P^{RX}_{(1,i,k_e,k_t)}}{\underbrace{\sum_{k' \neq k_e} P^{RX}_{(1,i,k',k_t)}}_{\substack{\text{помехи внутри} \\ \text{своего сектора}}} + \underbrace{\sum_{k_e} \sum_{i' \neq i} P^{RX}_{(1,i',k_e,k_t)}}_{\substack{\text{помехи внутри} \\ \text{своей соты}}} + \underbrace{\sum_{k_e} \sum_{i} \sum_{j \neq 1} P^{RX}_{(j,i',k_e,k_t)}}_{\substack{\text{помехи от} \\ \text{соседних сот}}} P_N}. \quad (14)$$

Рассмотрим слагаемые помех SNOI в знаменателе (14). Первое слагаемое определяет суммарный вклад помех от направленных радиолиний из набора $L_{(1,i,k')}, k' \neq k$ внутри своего сектора центральной соты базовой станции gNB_1. Второе слагаемое определяет суммарный вклад помех от направленных радиолиний из набора $L_{(1,i',k)}$ других секторов $\mathrm{s}_{i'}$, $i' \neq i$ своей соты базовой станции gNB_1. Третье слагаемое определяет суммарный вклад помех от направленных радиолиний из набора $L_{(j,i,k)}$ трех секторов s_i, окружающих сот базовых станций $\mathrm{gNB}_j, j = 2, \dots, 7$.

Из выражения (14) следует, что каждый луч является источником помех для других направленных радиолиний как в своем секторе соты, так и в других секторах других сот. В ИМ оценка SINR выполняется для трех сценариев с учетом помех: внутри своего сектора (Sector) – сценарий *S*, внутри своего сектора и своей соты (Cell) – сценарий *S* + *C* и внутри своего сектора, своей соты и окружающих сот модели сети (Network) – сценарий *S* + *C* + *N*.

Пропускную способность (ПС) в имитационной модели выполняется в точках $\mathbf{x}_k$ истинных местоположений пользовательских устройств UE_k из набора $k_t \in \mathbb{K}_t$ при ориентации трафиковых лучей на точки $\hat{\mathbf{x}}_k$ оценок координат UE_k из набора $k_e \in \mathbb{K}_e$. Нормированную ПС в канале «вниз» в бит/с/Гц для направленных радиолиний из набора $L_{(1,i,k)}$ в трех секторах центральной соты базовой станции gNB_1 можно оценить по формуле [18]:

$$T_{(1,i,k_e,k_t)} = \log_2\left(1 + SINR_{(1,i,k_e,k_t)}\right). \quad (15)$$

Метрика SINR может выступать косвенной характеристикой допустимого пространственного уплотнения одновременных передач и зависит от ряда факторов. В настоящей имитационной модели отношение сигнал/(шум + помеха) исследуется в зависимости от: 1) погрешности позиционирования σ; 2) радиуса соты *R*; 3) числа *K* пользовательских устройств в одном секторе.

Формализуем далее процедуры работы имитационной модели совокупности радиолиний с диаграммообразованием LAB.

2.4. Процедуры работы имитационной модели совокупности радиолиний с LAB

Рисунок 7 иллюстрирует последовательность процедур работы имитационной модели совокупности радиолиний с диаграммообразованием LAB для заданного сценария территориального развертывания.

Началом работы имитационной модели является процедура инициализации территориального плана гексагональной модели сверхплотной СРД, которая реализует территориальное развертывание семи базовых станций $\mathrm{gNB}_j, j = 1, \dots, 7$, каждая с тремя секторами s_i; внутри каждого сектора распределяются *K* устройств UE_k, в точках ОК $\hat{\mathbf{x}}_k$ из набора $k_e \in \mathbb{K}_e$. Число пользовательских устройств *K* в каждом секторе не может быть больше числа элементов АР N^2. В результате получается набор $L_{(j,i,k_e)}$. В ИМ реализовано равномерное распределение UE_k на площади каждого сектора s_i.

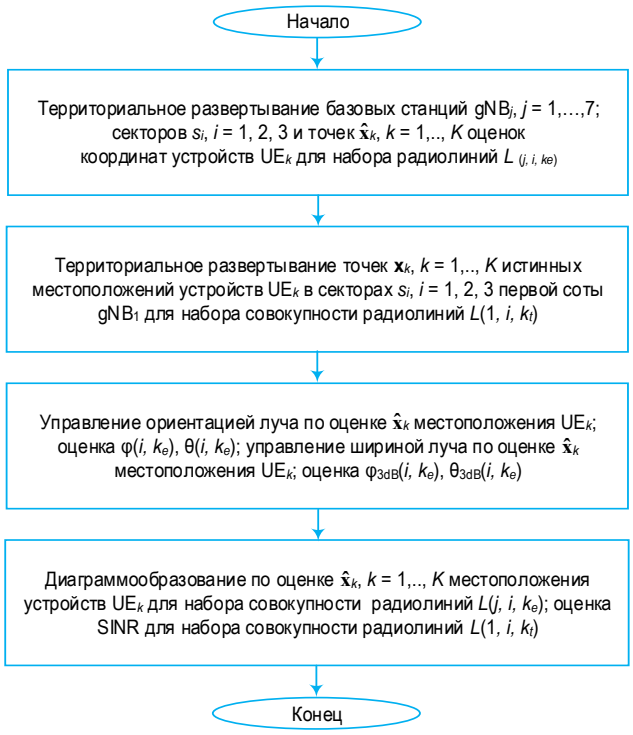


Рис. 7. Процедуры модели диаграммообразования LAB

Fig. 7. Location-Aware Beamforming Model Procedures

Далее в ИМ инициализируется сценарий территориального развертывания пользовательских устройств UE_k, в точках истинных местоположений $\mathbf{x}_k$ из набора $k_t \in \mathbb{K}_t$ в трех секторах центральной соты первой базовой станции gNB_1 (см. рисунок 1). Каждая точка $\mathbf{x}_k$ находится в круге неопределенности местоположения с центром в точке $\hat{\mathbf{x}}_k$ и радиусом $\sigma/2$ (см. рисунок 3). В результате получается набор $L_{(1,i,k_t)}$. Направленные радиолинии окружающих сот базовых станций $\mathrm{gNB}_j, j = 2, \ldots, 7$ служат для моделирования помех от соседних сот.

После инициализации территориального плана модули управления лучом реализуют процедуры вычисления ориентации и ширины лучей по точкам ОК $\hat{\mathbf{x}}_k$ из набора $k_e \in \mathbb{K}_e$. В результате получается совокупность $L_{(j,i,k_e)}$ направленных радиолиний, каждая из которых характеризуется кортежем из четырех углов $\varphi_{(i,k)}$, $\theta_{(i,k)}$, $\varphi_{3\mathrm{dB}(i,k)}$, $\theta_{3\mathrm{dB}(i,k)}$.

После настройки ориентации и ширины луча в модели выполняется формирование направленных радиолиний и оценки бюджета направленных радиолиний для полезного сигнала SOI и помех SNOI.

В заключении имитационная модель реализует вычисление отношения сигнал/(шум + помеха) $SINR_{(1,i,k_e,k_t)}$ по формуле (14) для точек $\mathbf{x}_k$ истинных местоположений пользовательских устройств UE_k из набора $k_t \in \mathbb{K}_t$ при ориентации трафиковых лучей на точки $\hat{\mathbf{x}}_k$ ОК UE_k из набора $k_e \in \mathbb{K}_e$. Также в трех секторах центральной соты gNB_1 рассчитывается набор значений $SINR_{(1,i,k_e,k_e)}$ для точек $\hat{\mathbf{x}}_k$ оценок координат пользовательских устройств UE_k из набора $k_e \in \mathbb{K}_e$ при ориентации лучей на точки $\hat{\mathbf{x}}_k$ из набора $k_e \in \mathbb{K}_e$. Дополнительно может быть выполнена оценка нормированной ПС $T_{(i,k_e,k_t)}$ по (15).

Формализуем далее параметры сценария имитационной модели совокупности радиолиний с диаграммообразованием LAB.

2.5. Сценарий имитационного моделирования совокупности радиолиний с LAB

Таблица 2 содержит параметры сценария ИМ совокупности радиолиний с LAB.

ТАБЛИЦА 2. Параметры сценария ИМ

TABLE 2. Simulation Model Scenario Parameters

Символ	Описание	Значение
–	территориальный план gNB	гексагональная сетка
J	число базовых станций gNB	7
I	число секторов на каждой gNB	3 (с ориентацией по азимуту 120°)
K	максимальное число одновременно работающих устройств в каждом секторе	64
R	максимальный размер соты	100 м
N^2	число элементов АР в каждом секторе	32×32
$\lvert\mathbb{I}\rvert$	общее число секторов	$J \cdot I = 21$
$\lvert\mathbb{K}_e\rvert$	максимальное число радиолиний SNOI из набора $k_e \in \mathbb{K}_e$	$J \cdot I \cdot K = 1344$
$\lvert\mathbb{K}_t\rvert$	максимальное число радиолиний SOI из набора $k_t \in \mathbb{K}_t$	$1 \cdot I \cdot K = 192$
–	распределение точек $\hat{\mathrm{x}}_k$ ОК UE_k в каждом секторе	случайное равномерное расположение на площади каждого сектора s_i
–	распределение точек x_k истинных местоположений UE_k в каждом секторе	случайное расположение в окружности с центром $\hat{\mathrm{x}}_k$ и диаметром σ в прямоугольной системе координат
P_s^{TX}	максимальная мощность передачи АР каждого сектора	40 дБм
$P_s^{\max}$	максимальная мощность передачи на один элемент АР	P_s^{TX}/N^2; равномерное распределение мощности между элементами АР
$G^{\max}$	максимальный КУ	15 дБи [29]
$A_{\min}^{AZ}$	максимальный коэффициент подавления задних лепестков	25 дБ [18]
$A_{\min}^{EL}$	максимальный коэффициент подавления боковых лепестков	20 дБ [18]
B_S	ширина полосы частот сектора	80 МГц
f_S	несущая частота сектора	30 ГГц
h_{gNB}	высота подъема АР	15 м
h_{UE}	высота подъема точки $\hat{\mathrm{x}}_k$	1,5 м
$L_{(i,k_t)}^{PL}$	потери РРВ в пространстве в радиолинии (i, k_t)	3GPP Umi-Street Canyon LOS [31]

Символ	Описание	Значение
G_S^{TX}	КУ одного элемента АР передатчика	3 дБи [16]
G_S^{BF}	максимальный КУ АР при диаграммообразовании	$10\log_{10}(N^2)$ [16]
P_N	мощность шума	согласно ITU-R M.2412-0 [21] с коэффициентом шума 5 дБ и шириной полосы B_S
φ_{3dBmin}	минимальная ширина луча в горизонтальной плоскости (по азимуту)	3 °
θ_{3dBmin}	минимальная ширина луча в вертикальной плоскости (по углу места)	3 °
$\sigma/2$	неопределенность местоположения	{1 ... 10} м согласно PSL в 3GPP TS 22.261 [25]

Рисунок 8 иллюстрирует сценарий территориального распределения базовых станций $\text{gNB}_j, j = 1, \ldots, 7$, секторов s_i, точек $\hat{\mathbf{x}}_k$ оценок координат UE_k в каждом секторе s_i каждой базовой станции gNB_j (синие точки) и распределение точек $\mathbf{x}_k$ истинных местоположений UE_k в каждом секторе центральной соты первой базовой станции gNB_1 (красные точки).

Территориальное распределение набора из семи $J = 7$ базовых станций gNB на плоскости выполняется по модели гексагональной решетки [5, 6]; особенностью правильного шестиугольника является равенство его стороны R и радиуса описанной окружности. Дальность радиопокрытия каждой gNB моделируется параметром стороны правильного шестиугольника $R = 100$ м; такая дальность соответствует сценарию сверхплотной СРД 5G в городе [21]. Каждая базовая станция gNB образована тремя $I = 3$ неперекрывающимися в горизонтальной плоскости секторами. Общее число секторов в модели СРД 5G равно $|\mathbb{I}| = J \cdot I = 21$. Каждый сектор s_i оборудован АР из $N^2 = 64$ излучающих элементов.

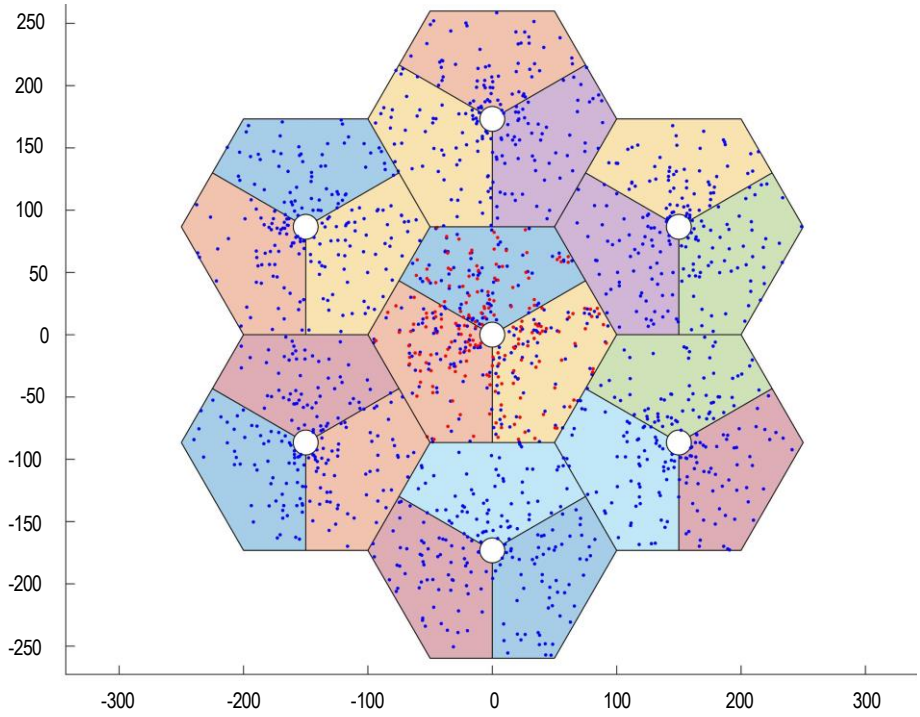


Рис. 8. Сценарий территориального распределения модели совокупности радиолиний

Fig. 8. Terrestrial Scenario of the Model of a Set of Radio Links

Для моделирования территориального распределения одновременно работающих в каждом секторе пользовательских устройств UE_k делается допущение о том, что каждый сектор s_i одновременно может обслужить максимальное их число K, которое ограничено сверху величиной N^2. Будем далее полагать, что общее число направленных радиолиний SNOI по всем секторам модели сети равно $|\mathbb{K}_e| = J \cdot I \cdot K = 1344$. Набор радиолиний формируется в каждом секторе s_i случайным образом в полярных координатах с центром в точке центра сектора s_i.

Оценка SINR выполняется в центральной соте базовой станции gNB_1, где истинные местоположения $\mathbf{x}_k$ пользовательских устройств UE_k формируются случайным образом в окружности с центром $\hat{\mathbf{x}}_k$ и диаметром σ в прямоугольной системе координат. Таким образом, отношение сигнал/(шум + помехи) оценивается для $|\mathbb{K}_t| = 192$ радиолиний SOI по фактическим местоположениям UE_k.

Рассмотрим параметры сценария ИМ (см. таблицу 2), определяющие оценку SINR. В каждом секторе моделируется передача на несущей частоте $f_S = 30$ ГГц с шириной полосы частот $B_S = 80$ МГц.

Общая максимальная мощность передатчика $P_s^{TX} =$ $= 40$ Вт, подводимая к АР, равномерно распределяется между всеми ее N^2 излучающими элементами, поэтому мощность, приходящая на один элемент, определяется как $P_s^{\max} = P_s^{TX}/N^2$.

В качестве модели PPB используется специфицированный в 3GPP TR 38.901 [31] сценарий Umi-Street Canyon LOS/NLOS с наличием LOS. Согласно данному сценарию, высота подвеса АР равна $h_{gNB} = 15$ м с креплением, например, на столбе, а высота UE $h_{UE} = 1{,}5$ м.

При настройке ориентации и ширины луча в имитационной модели используется численное ограничение на ширину диаграммы направленности антенной решетки в горизонтальной $\varphi_{3\text{dBmin}}$ и вертикальной $\theta_{3\text{dBmin}}$ плоскостях.

Технологические ограничения определяются методом диаграммообразования и конструктивным исполнением АР [29]. В настоящей имитационной модели принимаем минимальную ширину луча:

$$\varphi_{3\text{dBmin}} = \theta_{3\text{dBmin}} = 3\,^\circ.$$

3. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ МОДЕЛИ СОВОКУПНОСТИ РАДИОЛИНИЙ С LAB

Программная реализация имитационной модели доступна по ссылке [32] и включает следующие отдельные программные модули.

3.1. Инициализация параметров имитационной модели

Скрипт 2 содержит команды инициализации параметров имитационной модели; все параметры (см. таблицу 2) записываются в структуру udn.

Скрипт 2. Инициализация параметров имитационной модели

```
udn.cell_num=7;   % число сот
udn.sector_num=3; % число секторов
udn.UE_num=64;    % число устройств на сектор соты
udn.rcell=100;    % радиус соты, м
udn.accuracy=10;  % диаметр зоны местоположения UE, м
udn.radius=10;    % зона ограничения вокруг соты
udn.UE_h = 1.5;   % высота антенны UE, м
udn.gNB_h = 15;   % высота антенны gNB, м
udn.eff_h=udn.gNB_h-udn.UE_h; % эффективная высота, м
udn.txPowerDBm = 40; % мощность передачи, дБм
udn.txPower=(10.^((udn.txPowerDBm-30)/10)); % дБм→Вт
udn.Am = 25; % коэфф. подавления задних лепестков, дБ
udn.SLAv=20; % предельный уровень бок. лепестков, дБ
udn.GdB = 15;% коэфф. усиления АР малой соты, дБи
udn.G = 10^(udn.GdB/10); % КУ АР малой соты, раз
udn.Gtx=3;   % коэффициент усиления элемента АР, дБи
udn.fc=30 ;  % несущая частота, ГГц
udn.angle_min=3; % минимальное значение hpbw, градусы
udn.bw = 80e6;          % ширина полосы 80 МГц
udn.rxNoiseFigure = 5; % коэфф. шума приемника UE, дБ
udn.rxNoisePowerdB = ... % мощность шума, дБ, Вт
    -174 + 10*log10(udn.bw) + udn.rxNoiseFigure - 30;
udn.rxNoisePower = 10^(udn.rxNoisePowerdB/10);
udn.nrow = 32; % число элементов в строке АР
udn.ncol = 32; % число элементов в столбце АР
 % мощность передачи одного элемента АР, Вт
udn.txPowerSE = udn.txPower/(udn.nrow*udn.ncol);
udn.Gbf=10*log10(udn.nrow*udn.ncol);  % макс. КУ, дБ
```

3.2. Территориальное распределение базовых станций и секторов

Скрипт 3 содержит функцию lab_grid, реализующую гексагональный сценарий территориального распределения 7-ми сот, обслуживаемых трехсекторными базовыми станциями. Таблица 3 содержит формат и описание входных/выходных параметров функции lab_grid.

Скрипт 3. Территориальное распределение базовых станций и секторов

```
function [gNB, gNB_cell, gNB_sector]=lab_grid(r)
radius = 1; % область, недоступная для UE, м
% формирование гексагональной сетки с центральной
% сотой gNB в точке (0,0); центры соседних 6 сот
% определяются относительно центральной соты
gNB=[0 0]; % расположение центральной базовой станции
for theta=30:60:330
    x= r*sqrt(3)*cosd(theta);
    y= r*sqrt(3)*sind(theta);
    gNB = [gNB; x y];
end
% границы всех сот базовых станций
for i=1:length(gNB)
    xc = gNB(i,1);
    yc = gNB(i,2);
    xn = [(r+xc) (r/2+xc) (-r/2+xc) ...
          (-r+xc) (-r/2+xc) (r/2+xc) (r+xc)];
    yn = [yc (r*sqrt(3)/2+yc) (r*sqrt(3)/2+yc) ...
        yc (-r*sqrt(3)/2+yc) (-r*sqrt(3)/2+yc) yc];
    gNB_cell{i}=polyshape(xn,yn);
end
% границы секторов всех сот базовых станций
for i=1:length(gNB)
    xc = gNB(i,1);
    yc = gNB(i,2);
    % сектор 1
    x1 = [xc+3*r/4 xc+r/2 xc-r/2 xc-3*r/4 xc];
    y1 = [yc+r*sqrt(3)/4 yc+r*sqrt(3)/2 ...
          yc+r*sqrt(3)/2 yc+r*sqrt(3)/4 yc];
    % сектор 2
    x2 =[xc-3*r/4 xc-r xc-r/2 xc xc];
    y2 =[yc+r*sqrt(3)/4 yc yc-r*sqrt(3)/2 ...
         yc-r*sqrt(3)/2 yc];
    % сектор 3
    x3 =[xc xc+r/2 xc+r xc+3*r/4 xc];
    y3 =[yc-r*sqrt(3)/2 yc-r*sqrt(3)/2 ...
         yc yc+r*sqrt(3)/4 yc];
    % область, недоступная для UE, исключается
    th = 0:pi/50:2*pi;
    xunit = radius * cos(th) + xc;
    yunit = radius * sin(th) + yc;
    poly0 = polyshape(xunit(1:end-1),yunit(1:end-1));
    % формирование полных секторов
    poly1 = polyshape(x1,y1);
    poly2 = polyshape(x2,y2);
    poly3 = polyshape(x3,y3);
    % используемая область секторов
    gNB_sector{i,1} = subtract(poly1, poly0);
    gNB_sector{i,2} = subtract(poly2, poly0);
    gNB_sector{i,3} = subtract(poly3, poly0);
end
end
```

Сначала инициализируются координаты $\mathbf{x}_{gNB_1} =$ $= \left(x_{gNB_1}, y_{gNB_1}\right) = (0{,}0)$ базовой станции gNB_1, обслуживающей центральную соту. Координаты $\mathbf{x}_{gNB_j} = \left(x_{gNB_j}, y_{gNB_j}\right)$ остальных базовых станций $\text{gNB}_j, j = 2, \ldots, 7$ формируются относительно gNB_1.

Затем в цикле по числу базовых станций модели сети формируются границы сот в форме правильных шестиугольников. Форматом представления границ сот является массив gNB_cell ячеек размерности 1×7, каждая из которых содержит описание многоугольника в виде объекта polyshape среды Matlab [33]. Функция polyshape формирует многоугольник, задаваемый вершинами на плоскости, и возвращает объект polyshape, который характеризуется вершинами, сплошными областями и отверстиями. Также в цикле по числу базовых станций формируются границы секторов в формате объектов polyshape.

ТАБЛИЦА 3. Параметры функции lab_grid

TABLE 3. Function lab_grid Parameters

Параметр	Формат	Описание
Входные параметры		
r	скаляр	радиус соты
Выходные параметры		
gNB	матрица 7×2	матрица координат базовых станций на плоскости
gNB_cell	массив ячеек 1×7	массив ячеек границ сот в формате координат вершин правильного шестиугольника на плоскости
gNB_sector	массив ячеек 7×3	массив ячеек границ секторов в формате координат вершин многоугольника сектора на плоскости

В имитационной модели делается допущение о том, что в области poly0 непосредственной близости от расположения базовой станции, определяемой параметром radius, нахождение пользовательских устройств UE исключено, поэтому границы секторов gNB_sector формируются из границ полных секторов poly1, poly2 и poly3, за исключением областей poly0 с использованием функции subtract среды Matlab [34] (см. скрипт 3).

3.3. Территориальное распределение пользовательских устройств

Скрипт 4 содержит функцию lab_deploy, реализующую территориальное распределение UE внутри сот и секторов сформированной ранее модели сверхплотной сети. Таблица 4 содержит формат и описание выходных параметров функции lab_deploy.

ТАБЛИЦА 4. Выходные параметры функции lab_deploy

TABLE 4. Function **lab_deploy** *Output Parameters*

Параметр	Формат	Описание
UE_est	массив ячеек 7×3	массив ячеек ОК пользовательских устройств по набору сот и секторов
UE_true	массив ячеек 1×3	массив ячеек координат истинных местоположений UE в наборе секторов первой соты

Скрипт 4. Территориальное распределение UE

```
function [UE_est, UE_true]= ...
                lab_deploy(udn, gNB, gNB_sector)
UE_est{udn.cell_num,udn.sector_num}=[];
UE_true{1,udn.sector_num}=[];
for j=1:udn.cell_num % цикл по числу сот
    for i=1:udn.sector_num % цикл по числу секторов
        for k=1:udn.UE_num % цикл по числу устройств
            if i==1 % сектор 1
                theta=30:0.1:150;
            elseif i==2 % сектор 2
                theta=150:0.1:270;
            elseif i==3 % сектор 3
                theta=-90:0.1:30;
            end
            % для радиолиний UE_est во всех сотах
            c1=0;
            while c1<1
                key1 = randi([1, length(theta)]);
                % выбор r1 между 0-90 + udn.radius
                r1=((udn.rcell-udn.radius) ...
                        *rand(1,1))+udn.radius;
                theta1=theta(key1);
                % преобразование в прямоугольную СК
                x1=gNB(j,1)+r1*cosd(theta1);
                y1=gNB(j,2)+r1*sind(theta1);
                xy1=[x1,y1];
                % проверка нахождения в секторе
                if isinterior(gNB_sector{j,i},xy1)==1
                    c1=1;
                    UE_est1 = xy1;
                end
            end % while c1<1
            UE_est{j,i}=[UE_est{j,i}; UE_est1];
            if j==1 % для радиолиний gNB_UE_link_true
                c2=0;
                while c2<1
                    % для равномерного распределения
                    ru2 = (udn.accuracy/2) ...
                        *sqrt(rand(1,1));
                    theta2 = 2*pi*rand(1,1);
                    x2=UE_est1(1)+ru2*cos(theta2);
                    y2=UE_est1(2)+ru2*sin(theta2);
                    xy2=[x2,y2];
                    if isinterior(...
                        gNB_sector{1,i},xy2)==1
                        c2=1;
                        UE_true2 = xy2;
                    end
                end % while c2<1
                UE_true{j,i}=[UE_true{j,i};UE_true2];
            end % if j==1
        end % цикл по числу устройств
    end % цикл по числу секторов
end % цикл по числу сот
```

3.4. Настройка ориентации луча по местоположению устройства

Скрипт 5 содержит функцию lab_link, реализующую настройку ориентации луча по азимуту и углу места в направленных радиолиниях по ОК и истинному местоположению UE. Таблица 5 содержит формат и описание выходных параметров функции lab_link.

Скрипт 5. Настройка ориентации луча по местоположению устройства

```
function [az_est, el_est, az_tru, el_tru] = ...
    lab_link(udn, gNB, UE_est, UE_tru)
az_est{udn.cell_num,udn.sector_num}=[];
el_est{udn.cell_num,udn.sector_num}=[];
```

```
az_tru{1,udn.sector_num}=[];
el_tru{1,udn.sector_num}=[];
for j=1:udn.cell_num % цикл по числу сот
    for i=1:udn.sector_num % цикл по числу секторов
        for k=1:udn.UE_num % цикл по числу устройств
            UE_loc_est = UE_est{j,i}(k,:);
            gNB_loc=[gNB(j,1),gNB(j,2)];
            dist2D_loc_est=norm(UE_loc_est-gNB_loc);
            % для gNB_UE_loc_est
            % угол ориентации по азимуту
            steer_loc_est = ...
                  evalsteer(i,UE_loc_est,gNB_loc);
            % угол наклона
            tilt_loc_est = ...
                atan2d(udn.eff_h, dist2D_loc_est);
            % заполнение массивов ориентаций луча
            az_est{j,i}=[az_est{j,i}; steer_loc_est];
            el_est{j,i}=[el_est{j,i}; tilt_loc_est];
            if j==1 % для gNB_UE_loc_tru
                UE_loc_tru = UE_tru{j,i}(k,:);
                dist2D_loc_tru = ...
                norm(UE_loc_tru-gNB_loc);
                % угол ориентации по азимуту
                steer_loc_true = ...
                   evalsteer(i,UE_loc_tru,gNB_loc);
                % угол наклона
                tilt_loc_true = ...
                   atan2d(udn.eff_h, dist2D_loc_tru);
                % заполнение массивов ориентаций луча
                az_tru{j,i} = ...
                  [az_tru{j,i}; steer_loc_true];
                el_tru{j,i} = ...
                  [el_tru{j,i}; tilt_loc_true];
            end % if j==1
        end % цикл по числу устройств
    end % цикл по числу секторов
end % цикл по числу сот
end
```

ТАБЛИЦА 5. Выходные параметры функции `lab_link`

TABLE 5. Function **`lab_link`** *Output Parameters*

Параметр	Формат	Описание
`az_est`	массив ячеек 7×3	массив ячеек ориентации направленных радиолиний по азимуту для ОК пользовательских устройств по набору сот и секторов
`el_est`	массив ячеек 7×3	массив ячеек ориентации направленных радиолиний по углу места для ОК пользовательских устройств по набору сот и секторов
`az_tru`	массив ячеек 1×3	массив ячеек ориентации направленных радиолиний по азимуту для координат истинных местоположений UE в наборе секторов первой соты
`el_tru`	массив ячеек 1×3	массив ячеек ориентации направленных радиолиний по углу места для координат истинных местоположений UE в наборе секторов первой соты

3.5. Настройка ширины луча по местоположению устройства

Скрипт 6 содержит функцию `lab_hpbw`, реализующую управление шириной луча в горизонтальной и вертикальной плоскостях для направленных радиолиний по ОК пользовательских устройств. Таблица 6 содержит формат и описание выходных параметров функции `lab_hpbw`.

Скрипт 6. Настройка ширины луча по местоположению устройства

```
function [az_3dB, el_3dB] = ...
                lab_hpbw(udn, gNB_loc, UE_est)
% формирование массивов ширины луча
% по азимуту и углу места в радиолиниях UE_est
az_3dB{udn.cell_num,udn.sector_num}=[];
el_3dB{udn.cell_num,udn.sector_num}=[];
for j=1:udn.cell_num % цикл по числу сот
    for i=1:udn.sector_num % цикл по числу секторов
        for k=1:udn.UE_num % цикл по числу устройств
            gNB_locj=[gNB_loc(j,1), gNB_loc(j,2)];
            UE_loc_est = [UE_est{j,i}(k,1), ...
                                  UE_est{j,i}(k,2)];
            dist2D_est=norm(UE_loc_est-gNB_locj);
            rc=udn.accuracy/2; % радиус окружности
            % вычисление az3dB в зоне местоположения
            % точки пересечения двух окружностей
            [xout,yout] = circcirc(gNB_locj(1,1), ...
                gNB_locj(1,2), dist2D_est, ...
                UE_loc_est(1), UE_loc_est(2),rc);
            p1=[xout(1,1),yout(1,1)];
            p2=[xout(1,2),yout(1,2)];
            % расстояние точки-центр
            dp1=norm(gNB_locj-p1);
            dp2=norm(gNB_locj-p2);
            dpp=norm(p1-p2);
            % эффективное расстояние от центра
            dp1_eff=sqrt((dp1^2)+(udn.eff_h^2));
            dp2_eff=sqrt((dp2^2)+(udn.eff_h^2));
            az3dB=acosd(((dp1_eff^2)+(dp2_eff^2)- ...
                (dpp^2))/(2*dp1_eff*dp2_eff));
            if az3dB < udn.angle_min
                az3dB = udn.angle_min;
            end
            % вычисление el3dB в зоне местоположения
            theta = 0 : 0.01 : 2*pi;
            xc = UE_loc_est(1) + rc*cos(theta);
            yc = UE_loc_est(2) + rc*sin(theta);
            % самая ближняя/дальняя точка от gNB
            near=dist2D_est;
            far=near;
            for n=1:length(xc)
                point=[xc(1,n) yc(1,n)];
                tmp=norm(gNB_locj-point);
                if tmp<near
                    near=tmp;
                    nearestp=point;
                end
                if tmp>far
                    far=tmp;
                    farthest=point;
                end
            end
            dpp=norm(farthest-nearestp);
            near_eff=sqrt((near^2)+(udn.eff_h^2));
            far_eff=sqrt((far^2)+(udn.eff_h^2));
          el3dB = acosd(((near_eff^2)+ ...
           (far_eff^2)-(dpp^2))/(2*near_eff*far_eff));
            if el3dB < udn.angle_min
                el3dB = udn.angle_min;
            end
            az_3dB{j,i}=[az_3dB{j,i};az3dB];
            el_3dB{j,i}=[el_3dB{j,i};el3dB];
        end % цикл по числу устройств
    end % цикл по числу секторов
end % цикл по числу сот
end
```

ТАБЛИЦА 6. Выходные параметры функции `lab_hpbw`

TABLE 6. Function **`lab_hpbw`** *Output Parameters*

Параметр	Формат	Описание
`az_3dB`	массив ячеек 7×3	массив ячеек значений ширины луча направленных радиолиний в горизонтальной плоскости для ОК пользовательских устройств по набору сот и секторов
`el_3dB`	массив ячеек 7×3	массив ячеек значений ширины луча направленных радиолиний в вертикальной плоскости для ОК пользовательских устройств по набору сот и секторов

3.6. Оценка SINR по совокупности направленных радиолиний

Программная реализация процедур функции `lab_sinr`, реализующей оценку отношения SINR по совокупности радиолиний SOI и SNOI для направленных радиолиний по ОК и истинным местоположениям UE, доступна по ссылке [32]. Таблица 7 содержит формат и описание выходных параметров из ячеек размера 1×3 функции `lab_sinr`.

Функция `lab_sinr` содержит программные модули, реализующие оценку SINR по формуле (14) и включает расчет мощности: 1) сигнала SOI в радиолиниях секторов первой соты; 2) помех SNOI от радиолиний внутри своего сектора; 3) помех SNOI от радиолиний других секторов внутри своей соты; 4) помех SNOI от радиолиний других сот сети.

ТАБЛИЦА 7. Выходные параметры функции `lab_sinr`

TABLE 7. Function **`lab_sinr`** *Output Parameters*

Параметр	Массив значений SINR для направленных радиолиний
`SINR_S_est`	по ОК пользовательских устройств в наборе секторов первой соты, учитывающих помехи только в своем секторе
`SINR_S_tru`	по истинным местоположениям UE в наборе секторов первой соты, учитывающих помехи только в своем секторе
`SINR_SC_est`	по ОК пользовательских устройств в наборе секторов первой соты, учитывающих помехи в своем секторе и в двух соседних секторах данной соты
`SINR_SC_tru`	по истинным местоположениям UE в наборе секторов первой соты, учитывающих помехи в своем секторе и в двух соседних секторах данной соты
`SINR_SCN_est`	по ОК пользовательских устройств в наборе секторов первой соты, учитывающих помехи в своем секторе, в 2-х соседних секторах данной соты и 18-ти секторах окружающих шести сот
`SINR_SCN_tru`	по истинным местоположениям UE в наборе секторов первой соты, учитывающих помехи в своем секторе, в 2-х соседних секторах данной соты и 18-ти секторах окружающих шести сот

3.7. Функции оценки бюджета направленных радиолиний

3.7.1. Оценка диаграммы направленности АР

Скрипт 7 содержит программную реализацию процедуры оценки ДНА в горизонтальной и вертикальной плоскостях по формулам (10) и (11).

Скрипт 7. Оценка диаграммы направленности АР

```
function ARP = ...
        evalbarp(steer_True, steer_Est, az3dB, Am,...
                 tilt_True, tilt_Est, el3dB, SLAv)
% ДНА в горизонтальной плоскости
A_H = 12*(((steer_True-steer_Est)/az3dB).^2);
A_H=-(min(A_H,Am));
% ДНА в вертикальной плоскости
A_V=12*(((tilt_True-tilt_Est)/el3dB).^2);
A_V=-(min(A_V,SLAv));
ARP = A_H + A_V;     % совокупная ДНА
ARP(ARP<-Am) = - Am; % дБ
end
```

3.7.2. Оценка коэффициента усиления АР

Скрипт 8 содержит программную реализацию процедуры оценки КУ АР по формулам (7) и (8).

Скрипт 8. Оценка коэффициента усиления АР

```
function GAIN = ...
        evalgain(steer_True, steer_Est, az3dB, ...
                 tilt_True,  tilt_Est,  el3dB, Gbf)
GAIN = 10*log10(sinc(((...
% КУ в горизонтальной плоскости
steer_True - steer_Est)/(1.13*az3dB)).^2))...
    + 10*log10(sinc(((...
% КУ в вертикальной плоскости
    tilt_True - tilt_Est)/(1.13*el3dB)).^2))...
    + Gbf;
GAIN = real(GAIN);
end
```

3.7.3. Оценка потерь при РРВ

Скрипт 9 содержит программную реализацию процедуры оценки потерь при РРВ по формуле (12).

Скрипт 9. Оценка потерь при РРВ

```
function PL = evalfrisp(gNB_loc, UE_loc, eff_h, fc)
dist2D=norm(gNB_loc-UE_loc);       % 2D
dist3D=sqrt((eff_h)^2+(dist2D)^2); % 3D
% потери РРВ в радиолинии gNB_loc-UE_loc в LOS, дБ
PL = 32.4 + 21*log10(dist3D) + 20*log10(fc);
end
```

Далее представлены результаты моделирования совокупности радиолиний с диаграммообразованием LAB по набору сценариев.

4. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ СОВОКУПНОСТИ РАДИОЛИНИЙ С LAB

В настоящем разделе представлены результаты оценки SINR в модели совокупности радиолиний с диаграммообразованием LAB в зависимости от: 1) точности позиционирования σ; 2) радиуса соты R; 3) числа устройств в секторе K.

Оценка SINR выполняется для набора радиолиний: 1) $L_{(1,i,k_t)}$ по точкам истинных местоположений $\mathbf{x}_k$ (сценарий $\mathrm{UE_{tru}}$) из набора $k_t \in \mathbb{K}_t$; 2) $L_{(j,i,k_e)}$

по точкам ОК $\hat{\mathbf{x}}_k$ (сценарий $\mathrm{UE_{est}}$) из набора $k_e \in \mathbb{K}_e$. Результирующее значение SINR усредняется по набору радиолиний трех секторов центральной соты $\mathrm{gNB_1}$. Оценка усредненного SINR приводится для трех сценариев с учетом помех: 1) сценарий S; 2) сценарий $S + C$; 3) сценарий $S + C + N$. Таблица 2 содержит остальные параметры сценария модели.

4.1. Оценка помех в зависимости от точности позиционирования

Рисунок 9a иллюстрирует зависимость SINR от точности позиционирования σ для числа пользователей $K = 64$ и радиуса соты $R = 100$ м.

Анализ графиков позволяет сделать следующие выводы: 1) с уменьшением погрешности позиционирования σ с 10 до 1 м отношение SINR увеличивается примерно на 25 дБ; 2) для набора радиолиний $L_{(1,i,k_t)}$ по точкам истинных местоположений $\mathbf{x}_k$ (сценарий $\mathrm{UE_{tru}}$) отношение SINR ожидаемо всегда ниже отношения SINR для набора радиолиний по точкам ОК $\hat{\mathbf{x}}_k$ (сценарий $\mathrm{UE_{est}}$), так как ориентация луча в ИМ осуществляется по точкам $\hat{\mathbf{x}}_k$; 3) наибольшее отношение SINR ожидаемо наблюдается для сценарий S; наименьшее отношение SINR ожидаемо наблюдается для сценария $S + C + N$); разница между этими двумя сценариями составляет порядка 5 дБ; 4) разница в отношении SINR для трех сценариев S, $S + C$ и $S + C + N$ не зависит от погрешности σ; 5) повышение точности на 1 м увеличивает SINR на ~2-3 дБ.

Рисунок 9b иллюстрирует зависимость HPBW в горизонтальной φ_{3dB} и вертикальной θ_{3dB} плоскостях от точности позиционирования σ. Анализ графиков позволяет сделать следующие выводы: 1) с уменьшением погрешности позиционирования σ с 10 до 1 м ширина луча в горизонтальной плоскости φ_{3dB} уменьшается с 14 ° до 3 °, а ширина луча в вертикальной плоскости θ_{3dB} уменьшается с 6 ° до 3 °; 2) требуемая ширина луча в вертикальной плоскости θ_{3dB} оказывается ниже требуемой ширины луча в горизонтальной плоскости φ_{3dB}.

4.2. Оценка помех в зависимости от размера соты

Рисунок 9c иллюстрирует зависимость SINR от размера соты R при точности позиционирования $\sigma = 3$ м для числа пользователей $K = 64$.

Анализ графиков при прочих равных условиях позволяет сделать следующие выводы: 1) с увеличением размера соты R от 20 до 300 м отношение SINR увеличивается примерно на 30 дБ для сценария S и сценария $S + C$); 2) для сценария $S + C + N$ отношение SINR с увеличением размера соты сначала увеличивается, а при достижении некоторого порогового размера $R > 150$ м начинает уменьшаться.

Это можно объяснить тем, что *адаптация ориентации и ширины луча в имитационной модели осуществляется по местоположению устройства внутри соты, размер которой варьируется*.

Ранее уже отмечалось, что чем меньше σ, тем у́же в горизонтальной и вертикальной плоскостях будет луч; и наоборот, чем больше σ, тем шире будет луч. Неопределенность местоположения в данном сценарии фиксирована и равна $\sigma = 3$ м, поэтому влияние на уровень помех должна оказывать удаленность пользовательского устройства.

Также ранее говорилось о том, что чем ближе UE_k располагается к обслуживающему сектору s_i, тем шире получится луч; и наоборот, чем дальше UE_k располагается от s_i, например, на границе обслуживающего сектора, тем у́же получится луч. При увеличении размера соты R и фиксированном ограничении на минимальную ширину луча в горизонтальной $\varphi_{\mathrm{3dBmin}}$ и вертикальной θ_{3dBmin} плоскостях может наступить *пороговая ситуация, когда адаптированный по местоположению* UE_k *луч окажется недостаточно узким для заданного удаления* UE_k от s_i. Данную гипотезу косвенно подтверждает характер зависимости ширины луча HPBW от размера соты R.

Рисунок 9d иллюстрирует зависимость ширины луча HPBW от размера соты R при точности позиционирования $\sigma = 3$ м. Анализ графиков позволяет сделать следующие выводы: 1) требуемая ширина луча в вертикальной плоскости θ_{3dB} оказывается ниже требуемой ширины луча в горизонтальной плоскости φ_{3dB}; 2) с увеличением размера соты R с 20 до 300 м ширина луча в горизонтальной плоскости φ_{3dB} уменьшается с 8 ° до 3,5 °, а ширина луча в вертикальной плоскости θ_{3dB} уменьшается с 6 ° до 3 °; 3) при увеличении размера соты R с 20 до 100 м ширина луча в горизонтальной φ_{3dB} и вертикальной θ_{3dB} плоскостях уменьшается достаточно быстро; после значения $R = 100$ м скорость убывания HPBW заметно уменьшается; 4) установившаяся фиксированная ширина луча в вертикальной плоскости θ_{3dB} после некоторого порогового размера соты $R > 150$ м позволяет *сформулировать гипотезу о том, что если бы ограничение на HPBW было бы меньше, то тенденция уменьшения* θ_{3dB} *могла бы сохраниться, а уровень SINR в сценарии* $S + C + N$ *повторял бы характер аналогичной зависимости в сценариях* S *и* $S + C$.

Подтверждением сформулированной гипотезы являются графики на рисунках 9e, 9f, построенные при уменьшении ограничения на допустимую ширину луча HPBW в горизонтальной $\varphi_{\mathrm{3dBmin}}$ и вертикальной θ_{3dBmin} плоскостях с 3 ° до 1 °.

Анализ графиков на рисунке 9e позволяет сделать вывод о том, что с увеличением размера соты R от 20 до 300 м отношение SINR увеличивается примерно на 35 дБ для всех сценариев учета помех. Анализ графиков на рисунке 9f показывает, что ширина луча в горизонтальной плоскости φ_{3dB} уменьшается с 8 ° до 2 °, а ширина луча в вертикальной плоскости θ_{3dB} уменьшается с 6 ° до 1 °.

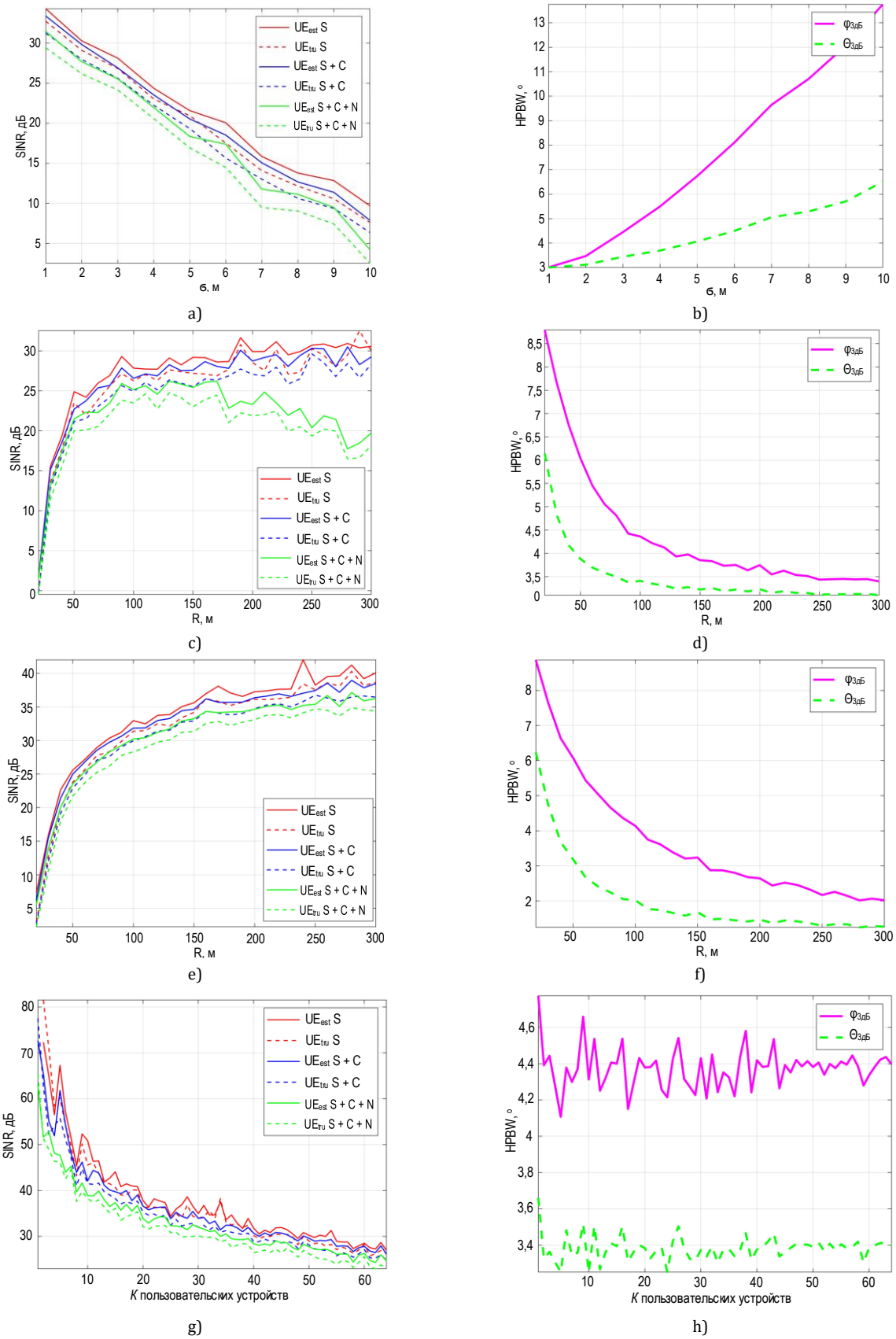


Рис. 9. Зависимость SINR (слева) и HPBW (справа) от точности позиционирования (a, b), размера соты (c, d), размера соты при $HPBW_{min}$ = 1 ° (e, f) и числа устройств в секторе (g, h)

Fig. 9. SINR (a) and HPBW (b) Dependence on UE Location Accuracy (a, b), Cell Size (c, d), ell Size with $HPBW_{min}$ = 1 ° (e, f), UE Number in Sector (g, h)

Сравнение графиков на рисунках 9c и 9e позволяет сделать вывод о том, что при увеличении R для сохранения характера зависимости SINR в сценарии $S + C + N$, аналогичного характеру зависимости SINR в сценариях S и $S + C$, допустимую ширину луча HPBW необходимо уменьшать с 3 ° до 1 °. Сравнение графиков на рисунках 9d и 9f позволяет сделать вывод о том, что с увеличением R разница между требуемой шириной луча в вертикальной θ_{3dB} и горизонтальной φ_{3dB} плоскостях уменьшается с 3 ° при малом R до 1° при большом R.

Например, при $R = 20$ м на площади сектора $S = \sqrt{3}/2\, R^2 \approx 346$ м^2 каждый из $K = 64$ UE занимает ~5 м^2, и для положительного SINR нужен луч в 1°.

4.3. Оценка помех в зависимости от числа устройств в секторе

Рисунок 9g иллюстрирует зависимость SINR от числа устройств в секторе K при размере соты $R = 100$ м и точности позиционирования $\sigma = 3$ м.

Анализ графиков (см. рисунки 9g, 9h) при прочих равных условиях позволяет сделать следующие выводы: 1) с увеличением числа пользовательских устройств K в секторе соты от 2 до 64 отношение SINR уменьшается для трех сценариев S, $S + C$ и $S + C + N$ примерно на 50 дБ ; 2) разница в отношении SINR для сценариев S, $S + C$ и $S + C + N$ уменьшается с увеличением числа пользовательских устройств K в секторе соты.

Рисунок 9h иллюстрирует зависимость ширины луча HPBW в горизонтальной φ_{3dB} и вертикальной θ_{3dB} плоскостях от числа устройств в секторе K при размере соты $R = 100$ м и точности позиционирования $\sigma = 3$ м. Анализ графиков позволяет сделать следующие выводы: 1) требуемая ширина луча в горизонтальной φ_{3dB} и вертикальной θ_{3dB} плоскостях не зависит от числа устройств в секторе K; 2) требуемая ширина луча в вертикальной θ_{3dB} плоскости ниже требуемой ширины луча в горизонтальной φ_{3dB} примерно на 1 °.

5. ВЫВОДЫ

В настоящей работе представлено описание разработанной и доступной для верификации имитационной модели совокупности направленных радиолиний, работающих по принципу LAB.

Теоретическая значимость разработанной модели заключается в установлении влияния ориентации и ширины луча базовой станции, а также погрешности определения местоположения пользовательского устройства на уровень пространственного уплотнения одновременных передач по критерию отношения сигнал/(шум + помеха).

Практическая значимость разработанной модели заключается в научном обосновании технических решений при построении и функционировании сверхплотных сетей радиодоступа диапазона миллиметровых волн с диаграммообразованием LAB пользовательских устройств.

Частными количественными результатами имитационного моделирования является установление зависимости отношения сигнал/(шум + помеха), а также требуемой ширины луча от точности позиционирования пользовательских устройств, размера соты и числа устройств в секторе.

С уменьшением погрешности позиционирования с 10 до 1 м отношение сигнал/(шум + помеха) увеличивается примерно на 25 дБ, а ширина луча в горизонтальной и вертикальной плоскостях уменьшаются с 14 ° до 3 ° и с 6 ° до 3 °, соответственно.

С увеличением размера соты от 20 до 300 м отношение SINR увеличивается примерно на 30 дБ при ограничении на ширину луча в 3 ° и примерно на 35 дБ при ограничении на ширину луча в 1 °. В последнем случае ширина луча в горизонтальной плоскости φ_{3dB} уменьшается с 8 ° до 2 °, а ширина луча в вертикальной плоскости θ_{3dB} уменьшается с 6 ° до 1 °. Исследование при двух ограничениях на ширину луча показало необходимость сужать луч при увеличении размера соты.

С увеличением числа пользовательских устройств в секторе соты от 2 до 64 отношение SINR уменьшается примерно на 50 дБ. При этом требуемая ширина луча в горизонтальной и вертикальной плоскостях не зависит от числа пользовательских устройств в секторе соты.

Разработанная модель является *инструментом решения научной проблемы* диаграммообразования на основе позиционирования в сверхплотных сетях радиодоступа миллиметрового диапазона. Установленная для совокупности направленных радиолиний взаимозависимость параметров размера соты, числа устройств и погрешности их позиционирования служит для научного обоснования допустимого по критерию SINR пространственного мультиплексирования.

Список источников

1. Фокин Г.А. Диаграммообразование на основе позиционирования в сверхплотных сетях радиодоступа миллиметрового диапазона. Часть 1. Модель двух радиолиний // Труды учебных заведений связи. 2023. Т. 9. № 4. С. 44–63. DOI:10.31854/1813-324X-2023-9-4-44-63

2. Фокин Г.А. Концепция диаграммообразования на основе позиционирования в сетях 5G // Вестник связи. 2022. № 10. С. 1–7.

3. Фокин Г.А. Сценарии позиционирования в сетях 5G // Вестник связи. 2020. № 3. С. 13–21.

4. Фокин Г.А. Моделирование сверхплотных сетей радиодоступа 5G с диаграммообразованием // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 5. С. 4–21. DOI:10.36724/2072-8735-2021-15-5-4-21.

5. Фокин Г.А. Модели диаграммообразования в сверхплотных сетях радиодоступа 5G. Часть 1. Оценка помех // Первая миля. 2021. № 3(95). С. 66–73. DOI:10.22184/2070-8963.2021.95.3.66.73

6. Фокин Г.А. Модели диаграммообразования в сверхплотных сетях радиодоступа 5G. Часть 2. Оценка разноса устройств // Первая миля. 2021. № 4(96). С. 66–73. DOI:10.22184/2070-8963.2021.96.4.66.72

7. Фокин Г.А. Процедуры выравнивания лучей устройств 5G NR // Электросвязь. 2022. № 2. С. 26–31. DOI:10.34832/ELSV.2022.27.2.003

8. Фокин Г.А. Модели управления лучом в сетях 5G NR. Часть 1. Выравнивание лучей при установлении соединения // Первая миля. 2022. № 1(101). С. 42–49. DOI:10.22184/2070-8963.2022.101.1.42.49

9. Фокин Г. Модели управления лучом в сетях 5G NR. Часть 2. Выравнивание лучей при ведении радиосвязи // Первая миля. 2022. № 3(103). С. 62–69. DOI:10.22184/2070-8963.2022.103.3.62.68

10. Fazliu Z.L., Malandrino F., Chiasserini C.F., Nordio A. MmWave Beam Management in Urban Vehicular Networks // IEEE Systems Journal. 2021. Vol. 15. Iss. 2. PP. 2798–2809. DOI:10.1109/JSYST.2020.2996909

11. Andrews J.G., Zhang X., Durgin G.D., Gupta A.K. Are we approaching the fundamental limits of wireless network densification? // IEEE Communications Magazine. 2016. Vol. 54. Iss. 10. PP. 184–190. DOI:10.1109/MCOM.2016.7588290

12. Roh W., Seol J.-Y., Park J., Lee B., Lee J., Kim Y., et al. Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results // IEEE Communications Magazine. 2014. Vol. 52. Iss. 2. PP. 106–113. DOI:10.1109/MCOM.2014.6736750

13. Chiaraviglio L., Turco S., Bianchi G., Blefari-Melazzi N. "Cellular Network Densification Increases Radio-Frequency Pollution": True or False? // IEEE Transactions on Wireless Communications. 2022. Vol. 21. Iss. 4. PP. 2608–2622. DOI:10.1109/TWC.2021.3114198

14. Chiaraviglio L., Rossetti S., Saida S., Bartoletti S., Blefari-Melazzi N. "Pencil Beamforming Increases Human Exposure to ElectroMagnetic Fields": True or False? // IEEE Access. 2021. Vol. 9. PP. 25158–25171. DOI:10.1109/ACCESS.2021.3057237

15. Thors B., Furuskär A., Colombi D., Törnevik C. Time-Averaged Realistic Maximum Power Levels for the Assessment of Radio Frequency Exposure for 5G Radio Base Stations Using Massive MIMO // IEEE Access. 2017. Vol. 5. PP. 19711–19719. DOI:10.1109/ACCESS.2017.2753459

16. Awada A., Lobinger A., Enqvist A., Talukdar A., Viering I. A simplified deterministic channel model for user mobility investigations in 5G networks // Proceedings of the International Conference on Communications (ICC, Paris, France, 21–25 May 2017). IEEE, 2017. DOI:10.1109/ICC.2017.7997079

17. Ali A., Karabulut U., Awada A., Viering I., Tirkkonen O., Barreto A.N., et al. System Model for Average Downlink SINR in 5G Multi-Beam Networks // Proceedings of the 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC, Istanbul, Turkey, 08–11 September 2019). IEEE, 2019. PP. 1–6. DOI:10.1109/PIMRC.2019.8904367

18. Yu B., Yang L., Ishii H. Load Balancing With 3-D Beamforming in Macro-Assisted Small Cell Architecture // IEEE Transactions on Wireless Communications. 2016. Vol. 15. Iss. 8. PP. 5626–5636. DOI:10.1109/TWC.2016.2563430

19. Harada H., Prasad R. Simulation and Software Radio for Mobile Communications. Artech House, 2002. 448 p.

20. ITU-R M.2135-1 (12/2009) Guidelines for evaluation of radio interface technologies for IMT-Advanced.

21. ITU-R M.2412-0 (10/2017) Guidelines for evaluation of radio interface technologies for IMT-2020.

22. 3GPP TS 23.273 V18.2.0 (06/2023) 5G System (5GS) Location Services (LCS); Stage 2 (Release 18).

23. 3GPP TS 22.071 V17.0.0 (03/2022) Location Services (LCS); Service description; Stage 1 (Release 17).

24. 3GPP TS 23.032 V18.0.0 (06/2023) Universal Geographical Area Description (GAD) (Release 18).

25. 3GPP TS 22.261 V19.3.0 (06/2023) Service requirements for the 5G system; Stage 1 (Release 19).

26. Gross F. Smart Antennas for Wireless Communications: With MATLAB. McGraw-Hill Professional, 2005. 288 p.

27. Balanis C.A. Antenna theory: analysis and design. John Wiley & Sons, 2016. 1104 p.

28. Mailloux R.J. Phased Array Antenna Handbook. Artech House, 2017. 691 p.

29. Hamdy M.N. Beamformers Explained. URL: www.commscope.com/globalassets/digizuite/542044-beamformer-explained-wp-114491-en.pdf (дата обращения 18.10.2023)

30. HBR 3.5 GHz 8x8 MIMO Panel Antenna. URL: https://halberdbastion.com/products/antenna-catalogue/hbr-35-ghz-8x8-mimo-panel-antenna (дата обращения 18.10.2023)

31. 3GPP TR 38.901 V17.0.0 (03/2022) Study on channel model for frequencies from 0.5 to 100 GHz (Release 17).

32. Имитационная модель совокупности радиолиний с диаграммообразованием на основе позиционирования в сетях 5G // GitHub. URL: https://github.com/grihafokin/LAB_system_level_rus (дата обращения 18.10.2023)

33. polyshape. 2-D polygonal shapes // MathWorks. URL: https://www.mathworks.com/help/matlab/ref/polyshape.html (дата обращения 18.10.2023)

34. subtract. Difference of two polyshape objects // MathWorks. URL: https://www.mathworks.com/help/matlab/ref/polyshape.subtract.html (дата обращения 18.10.2023)

References

1. Fokin G. Location Aware Beamforming in Millimeter-Wave Band Ultra-Dense Radio Access Networks. Part 1. Model of Two Links. *Proceedings of Telecommun. Univ.* 2023;9(4):44–63. DOI:10.31854/1813-324X-2023-9-4-44-63

2. Fokin G.A. Concept of Location-Aware Beamforming in 5G Networks. *Vestnik Ssviazy*. 2022;10:1–7.

3. Fokin G.A. Scenarios for Positioning in 5G Networks. *Vestnik Ssviazy*. 2020;3:13–21.

4. Fokin G.A. Simulation of ultra dense 5G radio access networks with beamforming. *T-Comm*. 2021;15(5):4–21. DOI:10.36724/2072-8735-2021-15-5-4-21

5. Fokin G.A. Beamforming models in ultra-dense 5G radio access networks. Part 1: Interference evaluation. *First mile*. 2021;3(95):66–73. DOI:10.22184/2070-8963.2021.95.3.66.73
6. Fokin G.A. Beamforming models in ultra-dense 5G radio access networks. Part 2: Device separation evaluation. *First mile*. 2021;4(96):66–73. DOI:10.22184/2070-8963.2021.96.4.66.72
7. Fokin G.A. Beam alignment procedures for 5G NR devices. *Elektrosvyaz*. 2022;2:26–31. DOI:10.34832/ELSV.2022.27.2.003
8. Fokin G.A. Beam management models in 5G NR networks. Part 1. Beam alignment during link establishment. *First mile*. 2022;1(101):42–49. DOI:10.22184/2070-8963.2022.101.1.42.49
9. Fokin G.A. Beam management models in 5G NR networks. Part 2. Beam alignment during radio communication. *First mile*. 2022;3(103):62–69. DOI:10.22184/2070-8963.2022.103.3.62.68
10. Fazliu Z.L., Malandrino F., Chiasserini C.F., Nordio A. MmWave Beam Management in Urban Vehicular Networks. *IEEE Systems Journal*. 2021;15(2):2798–2809. DOI:10.1109/JSYST.2020.2996909
11. Andrews J.G., Zhang X., Durgin G.D., Gupta A.K. Are we approaching the fundamental limits of wireless network densification? *IEEE Communications Magazine*. 2016;54(10):184–190. DOI:10.1109/MCOM.2016.7588290
12. Roh W., Seol J.-Y., Park J., Lee B., Lee J., Kim Y., et al. Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results. *IEEE Communications Magazine*. 2014;52(2):106–113. DOI:10.1109/MCOM.2014.6736750
13. Chiaraviglio L., Turco S., Bianchi G., Blefari-Melazzi N. "Cellular Network Densification Increases Radio-Frequency Pollution": True or False? *IEEE Transactions on Wireless Communications*. 2022;21(4):2608–2622. DOI:10.1109/TWC.2021.3114198
14. Chiaraviglio L., Rossetti S., Saida S., Bartoletti S., Blefari-Melazzi N. "Pencil Beamforming Increases Human Exposure to ElectroMagnetic Fields": True or False? *IEEE Access*. 2021;9:25158–25171. DOI:10.1109/ACCESS.2021.3057237
15. Thors B., Furuskär A., Colombi D., Törnevik C. Time-Averaged Realistic Maximum Power Levels for the Assessment of Radio Frequency Exposure for 5G Radio Base Stations Using Massive MIMO. *IEEE Access*. 2017;5:19711–19719. DOI:10.1109/ACCESS.2017.2753459
16. Awada A., Lobinger A., Enqvist A., Talukdar A., Viering I. A simplified deterministic channel model for user mobility investigations in 5G networks. *Proceedings of the International Conference on Communications, ICC, 21–25 May 2017, Paris, France*. IEEE; 2017. DOI:10.1109/ICC.2017.7997079
17. Ali A., Karabulut U., Awada A., Viering I., Tirkkonen O., Barreto A.N., et al. System Model for Average Downlink SINR in 5G Multi-Beam Networks. *Proceedings of the 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, 08–11 September 2019, Istanbul, Turkey*. IEEE; 2019. p.1–6. DOI:10.1109/PIMRC.2019.8904367
18. Yu B., Yang L., Ishii H. Load Balancing With 3-D Beamforming in Macro-Assisted Small Cell Architecture. *IEEE Transactions on Wireless Communications*. 2016;15(8):5626–5636. DOI:10.1109/TWC.2016.2563430
19. Harada H., Prasad R. Simulation and Software Radio for Mobile Communications. Artech House, 2002. 448 p.
20. ITU-R M.2135-1 (12/2009) *Guidelines for evaluation of radio interface technologies for IMT-Advanced*.
21. ITU-R M.2412-0 (10/2017) *Guidelines for evaluation of radio interface technologies for IMT-2020*.
22. 3GPP TS 23.273 V18.2.0 (06/2023) *5G System (5GS) Location Services (LCS); Stage 2 (Release 18)*.
23. 3GPP TS 22.071 V17.0.0 (03/2022) *Location Services (LCS); Service description; Stage 1 (Release 17)*.
24. 3GPP TS 23.032 V18.0.0 (06/2023) *Universal Geographical Area Description (GAD) (Release 18)*.
25. 3GPP TS 22.261 V19.3.0 (06/2023) *Service requirements for the 5G system; Stage 1 (Release 19)*.
26. Gross F. *Smart Antennas for Wireless Communications: With MATLAB*. McGraw-Hill Professional; 2005. 288 p.
27. Balanis C.A. *Antenna theory: analysis and design*. John Wiley & Sons; 2016. 1104 p.
28. Mailloux R.J. *Phased Array Antenna Handbook.* Artech House; 2017. 691 p.
29. Hamdy M.N. *Beamformers Explained*. URL: www.commscope.com/globalassets/digizuite/542044-beamformer-explained-wp-114491-en.pdf [Accessed 18.10.2023]
30. *HB Radiofrequency.* HBR 3.5 GHz 8x8 MIMO Panel Antenna. URL: https://halberdbastion.com/products/antenna-catalogue/hbr-35-ghz-8x8-mimo-panel-antenna [Accessed 18.10.2023]
31. 3GPP TR 38.901 V17.0.0 (03/2022) *Study on channel model for frequencies from 0.5 to 100 GHz (Release 17)*.
32. *GitHub*. A simulation model of a population of radio lines with diagramming based on positioning in 5G networks. URL: https://github.com/grihafokin/LAB_system_level_rus [Accessed 18.10.2023]
33. *MathWorks*. polyshape. 2-D polygonal shapes. URL: https://www.mathworks.com/help/matlab/ref/polyshape.html [Accessed 18.10.2023]
34. *MathWorks*. subtract. Difference of two polyshape objects. URL: https://www.mathworks.com/help/matlab/ref/polyshape.subtract.html [Accessed 18.10.2023]

Статья поступила в редакцию 14.10.2023; одобрена после рецензирования 29.10.2023; принята к публикации 02.11.2023.
The article was submitted 14.10.2023; approved after reviewing 29.10.2023; accepted for publication 02.11.2023.

Информация об авторе:

ФОКИН Григорий Алексеевич
доктор технических наук, доцент, профессор кафедры радиосвязи и вещания Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0002-5358-1895

DOI:10.31854/1813-324X-2023-9-5

Научная статья
УДК 004.942
DOI:10.31854/1813-324X-2023-9-5-66-78

Модель классификации трафика в программно-конфигурируемых сетях с элементами искусственного интеллекта

Василий Сергеевич Елагин, v.elagin@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация: *Классификация приложений необходима для повышения производительности сети. Однако при постоянном росте числа пользователей и приложений, а также масштабирования сетей, традиционные методы классификации не могут справляться в полной мере с идентификацией и классификацией сетевых приложений с необходимым уровнем задержки. Применение технологии глубокого обучения совместно с особенностями архитектуры программно-конфигурируемых сетей (SDN, аббр. от англ. Software-Defined Networking) позволит реализовать новую гибридную глубокую нейронную сеть для классификации приложений, которая сможет обеспечить высокую точность классификации без ручного выбора и извлечения признаков. В предлагаемой структуре предложена классификация приложений, с учетом логического централизованного управления на контроллере SDN. Обработанные данные используются для обучения гибридной глубокой нейронной сети, состоящей из многоуровневого автокодировщика, с высокой размерностью скрытого слоя и выходного слоя на базе регрессии softmax. Необходимые параметры сетевого потока могут быть получены при обработке трафика многоуровневым автокодировщиком вместо ручной обработки. Слой регрессии softmax используется в качестве конечного классификатора приложений. В статье приведены результаты моделирования, которые демонстрируют преимущества предложенного метода классификации, по сравнении с методом опорных векторов.*

Ключевые слова: *программно-конфигурируемые сети, ПКС, нейронная сеть, классификация трафика, регрессия softmax*

Финансирование: *научная статья подготовлена в рамках прикладных научный исследований СПбГУТ, регистрационный номер 123060900012-6 в ЕГИСУ НИОКТР.*

Ссылка для цитирования: Елагин В.С. Модель классификации трафика в программно-конфигурируемых сетях с элементами искусственного интеллекта // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 66–78. DOI:10.31854/1813-324X-2023-9-5-66-78

Traffic Classification Model in Software-Defined Networks with Artificial Intelligence Elements

Vasiliy Elagin, v.elagin@sut.ru

The Bonch-Bruevich Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Abstract: *Application classification is essential to improve network performance. However, with the constant growth in the number of users and applications, as well as the scaling of networks, traditional classification methods cannot fully cope with the identification and classification of network applications with the required level of delay. The use of deep learning technology together with the architecture features of software-defined networks (SDN) will allow the implementation of a new hybrid deep neural network for application classification, which can*

provide high classification accuracy without manual selection and feature extraction. The proposed structure proposes a classification of applications, taking into account the logical centralized management on the SDN controller. The processed data is used to train a hybrid deep neural network consisting of stacked autoencoder with a high dimensionality of the hidden layer and an output layer based on softmax regression. The necessary network flow parameters can be obtained by processing traffic with a stacked auto-encoder instead of manual processing. The softmax regression layer is used as the final application classifier. The article presents simulation results that demonstrate the advantages of the proposed classification method in comparison with the support vector machine.

Keywords*: software-defined networks, SDN, neural network, traffic classification, softmax regression*

Funding*:the article was prepared within the framework of applied scientific research of SPbSUT, registration number 123060900012-6 in the EGISU R&D.*

For citation: Elagin V. Traffic Classification Model in Software-Defined Networks with Artificial Intelligence Elements. *Proceedings of Telecommun. Univ.* 2023;9(5):66–78. DOI:10.31854/1813-324X-2023-9-5-66-78

Введение

Современные требования к обработке и управлению трафиком в сетях связи требуют постоянного увеличения производительности сетевых элементов и серверной инфраструктуры, однако, традиционная сетевая архитектура не удобна для сбора и обработки больших объемов данных и принятия решений из-за своей распределенности и децентрализованности. Поэтому целесообразно рассмотреть архитектуру программно-конфигурируемых сетей (SDN, *аббр. от англ.* Software-Defined Networking), в которых используется логически централизованное управление, в рамках которого можно организовать сбор нужных параметров. Кроме того, SDN на уровне управления (*англ.* Control Plane) открыты для развертывания новых сетевых сервисов. Архитектура SDN позволяет обеспечить полное управление трафиком (*англ.* Traffic Engineering) для повышения производительности сети и управления. Что еще более важно, SDN может значительно облегчить сбор и обработку больших объемов данных, благодаря централизованному управлению через SDN-контроллер [1, 2].

Максимально точная классификация сетевых приложений необходима для реализации различных сетевых функций, таких как динамическая маршрутизация трафика, трафик инжиниринг, обеспечение качества обслуживания (QoS, *аббр. от англ.* Quality of Service), прогнозирования событий и состояния сети [3].

Для повышения точности идентификации и классификации сетевых приложений за последнее время было предложено большое количество различных методов, в основном включая сигнатурный подход, основанный на заданных шаблонах трафика, подход, основанный на поведенческом анализе, и подход, основанный на машинном обучении [4]. В настоящее время активное развитие получила технология, основанная на машинном обучении и обработке больших данных, в которой обычно применяют нейронные сети с различными методами принятия решения: метод опорных векторов (SVM, *аббр. от англ.* Support Vector Machine), метод обратного распространения ошибки (*англ.* Backpropagation), Байесовские сети и алгоритм C4.5 для построения деревьев решений [5–13].

В отличие от традиционных методов сигнатурного анализа и поведенческого анализа, в методах классификации, основанных на машинном обучении, статистические характеристики трафика, включая размер пакета, время между пакетами, длительность сеанса и т. д., используются для идентификации и классификации сетевого приложения с помощью машинного обучения.

Нейронная сеть с небольшим числом слоев обладает ограниченной возможностью к обучению из-за небольшого количества слоев для нелинейного извлечения признаков и больше подходит для решения задач с ограниченным набором данных. Однако в связи с постоянным расширением масштабов сетей и наступлением эры больших данных сетевой трафик и число приложений стремительно растет. Это не позволяет традиционным нейронным сетям при растущем объеме сетевого трафика своевременно справляться с классификацией из-за ограниченной способности к обучению. Для извлечения дополнительных признаков из трафика, в том числе на прикладном уровне, на крупных сетях требуется нейронная сеть с глубоким обучением.

На данный момент активно используются модели глубоких нейронных сетей (ГНС), которые уже были апробированы для решения различных задач, в частности: сверточная нейронная сеть (CNN, *аббр. от англ.* Convolutional Neural Network), многослойный автокодировщик, глубокая сеть доверия (DBN, *аббр. от англ*. Deep Belief Network); они показали свою эффективность в областях распознавания, обработки речи, классификации признаков и объектов. Для обучения и дообучения ГНС обычно используются оригинальные алгоритмы – глубокого машинного обучения, в тех случаях, когда традиционные алгоритмы обучения не подходят по причине ограничения значений в областях локальных оптимумов и исчезающего градиента [14–15].

В данной статье будет предложен метод классификации сетевых приложений на основе гибридной модели глубокого обучения (DL – *аббр. от англ.* Deep Learning).

При этом необходимо отметить, что процесс классификации приложений определяется тремя необходимыми функциями: обучение модели, сам процесс классификации и последующая проверка и валидация результатов. При этом необходима постоянная обработка большого объема трафика и его параметров при предварительной обработке, построении обучающего и тестового наборов, разметки данных, классов и т. д.

Применение программно-конфигурируемых сетей SDN позволит упростить эти задачи благодаря декомпозиции функций управления и плоскости данных в SDN, а также наличию единого управляющего контроллера, который используется для виртуализации и быстрого внедрения новых сетевых функций, трафик инжиниринга и гибкого управления сетью и др. [16–22].

Эти возможности SDN могут значительно облегчить сбор и обработку больших массивов данных [23]. Используя логический централизованный контроллер и вычислительные возможности, можно значительно упростить сбор и обработку большого объема сетевого трафика.

Таким образом, целесообразно принять во внимание архитектуру SDN и технологию глубокого обучения для формирования нового метода классификации сетевых приложений, который будет состоять из следующих модулей:

– на основе архитектуры SDN предлагается новая модель классификации сетевых приложений с использованием глубокого обучения (структура позволяет легко собирать и обрабатывать большие массивы трафика для классификации приложений);

– комбинация многоуровневого автокодировщика и регрессионной модели softmax, позволяет разработать гибридную модель ГНС для классификации приложений; в этой модели многоуровневый автокодировщик используется для извлечения параметров потока и его свойств, а регрессия softmax используется для классификации сетевых приложений.

В дальнейшем в статье будут представлены результаты экспериментальной проверки предложенных моделей.

Существующие модели и решения для классификации трафика

Многоуровневый автокодировщик

Многоуровневый автокодировщик – это нейронная сеть с несколькими слоями Sparse Autoencoders. Структура автокодировщика состоит всего из 3-х типов слоев: входного слоя, скрытого слоя и слоя декодирования/реконструкции.

Стоит отметить основные отличия многоуровневого автокодировщика от других ГНС, например, ограниченной машины Больцмана (RBM, *аббр. от англ.* Restricted Boltzmann Machine) или сверточной нейронной сети:

1) в многоуровневом автокодировщике размерность ввода равна размерности вывода;

2) для обучения многоуровневого автокодировщика необходим только немаркированный (а не размеченный) набор данных, из-за неконтролируемого подхода к обучению (в отличие от CNN, представляющей собой алгоритм обучения с учителем, и DBN, состоящей из нескольких RBM и использующей алгоритм обучения с частичным привлечением учителя);

3) выходные результаты автокодировщика – это еще одно представление выходных данных, высокоуровневые характеристики потока могут быть извлечены посредством обучения многоуровневого автокодировщика алгоритмом без учителя.

Исходя из перечисленного выше, в данной статье был выбран многоуровневый автокодировщик в качестве нейронной сети с глубоким обучением для извлечения признаков потока.

Регрессия softmax

Модель регрессии softmax представляет собой поколение традиционной модели логистической регрессии и относится к алгоритму обучения с учителем. Он удобен при создании мультиклассовой логистической регрессии и заключается в том, чтобы не разбивать многоклассовые данные на несколько датасетов, используя бинарный классификатор, а сразу применять функции, которые позволяют работать с множеством классов. Это подходит для решения задачи многоклассовой классификации, поэтому имеет широкое применение во многих приложениях.

Для задачи классификации K классов вероятность того, что входные данные $x^{(i)}$ принадлежат классу j, может быть представлена как:

$$P\left(z^{(i)} = j \middle| x^{(i)}\right) = \frac{e^{\theta_j^T X(i)}}{\sum_{i=1}^{K} e^{\theta_i^T X(i)}}, \quad (1)$$

где $z^{(i)}$ – результат на выходе; θ_j – вектор входных параметров.

Кроме того, стоит отметить, что регрессия softmax часто используется в качестве выходного уровня гибридных ГНС как конечный классификатор [24]. В данной статье слой регрессии softmax планируется подключить к многоуровневому автокодировщику для повышения точности классификации приложений.

Стоит отметить, что модель регрессии softmax является одним из алгоритмов, реализующих множественную логистическую регрессию в случае категориальной переменной, поэтому применение данной регрессии позволяет исключить использование нескольких моделей множественной логистической регрессии для обработки большого числа классов трафика. Это, в основном, связано с практической реализацией моделей, так как большинство программных библиотек реализовано на низкоуровневых языках и оптимизированы под свои задачи, поэтому при программировании на более высоком уровне абстракции (как в нашем случае) возможны коллизии и неверные интерпретации отдельных классов или параметров. Поэтому для решения поставленной задачи многопараметрической классификации слой регрессии softmax планируется подключить к многоуровневому автокодировщику для повышения точности классификации приложений.

Большинство методов машинного обучения сосредоточены на эффективной настройке используемых методов и подбору комбинации признаков в потоке трафика, для повышения точности и полноты классификатора [4–13, 25, 26].

Перед обучением нейронной сети необходимо выбрать несколько параметров потока, так, например, в работе [4] выделено 37 параметров, включая наименование протокола, номера портов, продолжительность сессии, пропускную способность в пакетах и в байтах. Однако при экспоненциальном росте сетевого трафика в действующей сетевой архитектуре значительно усложняется задача сбора и обработки всего проходящего трафика, а также его обработка для получения искомых характеристик потока и повышения точности классификации

Существует новая архитектура выбора характеристик потока, которая может определить необходимое подмножество параметров для классификации различных типов сеансов связи, в рамках архитектуры SDN [26]. Данная архитектура состоит из менеджера потоков и системы выбора потока. Первый используется для исчисления характеристик потока, система выбора потока отвечает за анализ и интерпретацию этих характеристик.

В отдельных работах [11–13] были рассмотрены варианты применения архитектуры SDN для классификации сетевых приложений. Например, Прасад и Катаока [11] разработали многопутевой механизм пересылки пакетов с учетом приложений, объединив машинное обучение и SDN. Для получения данных о составе сервисов для построения классификатора приложений использовался алгоритм обучения деревьев решений C4.5, а в контроллер, как в единую точку управления и мониторинга сети, были интегрированы тренажер и классификатор, работающий на базе машинного обучения. Кроме того, в исследовании разработана новая архитектура применения машинного обучения для сбора данных и классификации трафика с учетом интеграции с SDN. В данной архитектуре контроллер получает статистику потоков от коммутаторов и агрегирует ее, в дальнейшем система применяет контролируемый алгоритм машинного обучения для классификации трафика [12].

Отдельные статьи предлагают применение классификации трафика для удовлетворения требований качества обслуживания QoS приложений, так, например, в работе [13] была разработана структура классификации трафика с учетом QoS в SDN. В этой структуре для обнаружения основных потоков применялась технология DPI (*от англ.* Deep Packet Inspection), а для классификации трафика с учетом QoS с помощью функции сопоставления использовался полууправляемый алгоритм машинного обучения. В частности, поток определенного приложения был сопоставлен с предварительно выделенным классом QoS в соответствии с его особенностями.

Хотя вышеупомянутые исследовательские работы учитывают архитектуру SDN для классификации приложений, эти методы в целом относятся к традиционной сети с неглубоким обучением, которая не может эффективно работать с массивными данными из-за ограниченных возможностей обучения функциям. В отличие от нее, сеть с глубоким обучением обладает более мощной способностью к изучению признаков и выделению параметров потока трафика [27].

Глубокая нейронная сеть в последнее время достаточно часто используется для обнаружения вторжений, прогнозирования трафика и классификации приложений [28–32]. Гибридные интеллектуальные архитектуры используются для различных целей, например, обнаружения сетевых аномалий [28, 29]. При этом исследователи предлагают различные варианты используемых нейронных сетей, в основном это RBM, иногда в комбинации с SVM.

Отдельно стоит отметить применение глубокого обучения для прогнозирования поведения сетевого трафика [30]. Для извлечения характеристик потока обычно используется модель традиционного многоуровневого автокодировщика или многоуровневого автокодировщика Левенберга – Марквардта [31]. В некоторых работах, например [32], сеть глубокого обучения для прогнозирования сетевого трафика предлагается комбинировать из DBN и слоя мультизадачной регрессии.

В данном исследовании предлагается применение архитектуры, объединяющей SDN и модели глубокого обучения. Контроллер SDN будет использоваться для обработки больших объемов

сетевого трафика и получения статистик потоков. Гибридную сетевую модель глубокого обучения, состоящую из многоуровневого автокодировщика и слоя регрессии softmax, планируется использовать для извлечения параметров потока и построения классификатора приложений.

Модельная сеть

Для идентификации потока предлагается ограничиться пятью основными параметрами: IP-адресом источника, IP-адресом назначения, транспортным протоколом (TCP или UDP), номером порта источника, номером порта назначения.

На рисунке 1 отражена схема предлагаемой структуры, в которой плоскость управления состоит из четырех основных функциональных модулей: мониторинга сети, сбора статистики потоков, обработки данных и глубокого обучения, а также базы данных, в которой хранится информация о параметрах потоков.

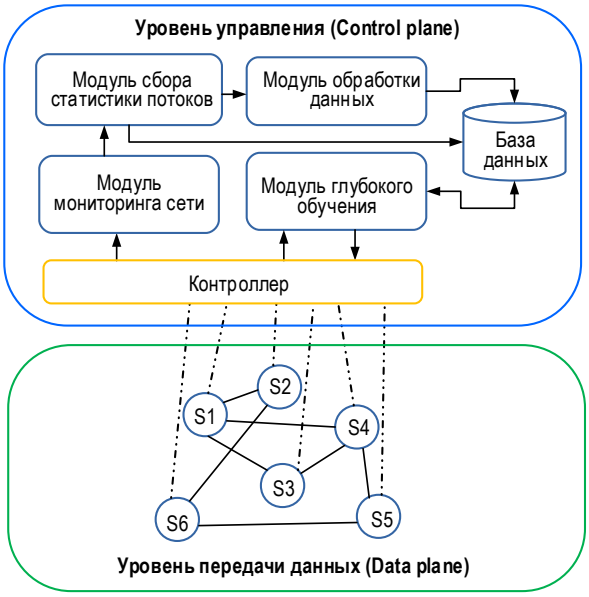


Рис. 1. Архитектура построения модельной сети для классификации приложений трафика

Fig. 1. The Application Classification Framework

В управлении сетевой инфраструктурой можно выделить ряд отдельных задач, которые должны решаться системой управления.

Модуль мониторинга сети – необходим для агрегации и анализа информации о сетевом трафике от коммутаторов в выделения информации о потоке (тип протокола, порт источника, порт назначения и т. д.).

Модуль сбора статистики потока – отвечает за обработку собранной информации о потоке и выделение из нее необходимых характеристик потока (например, продолжительность сессии, время между пакетами и распределение числа пакетов в потоке), а также данных о приложении (например, метка потока). В дальнейшем полученные параметры сохраняются в базе данных.

Модуль обработки данных – обрабатывает и использует характеристики потока и метки потока, хранящихся в базе данных, для создания обучающего набора в нейронной сети, а также тестового набора для встроенного классификатора приложений.

Модуль глубокого обучения – отвечает за обучение предложенной модели гибридной ГНС с использованием обучающего набора и тестового набора для работы встроенного классификатора приложений. По завершении обучения данный модуль при обнаружении нового потока классифицирует поступающий поток сетевых приложений в определенный класс приложений.

Информация, хранящаяся в базе данных, делится на 2 категории: информация из собранного сетевого трафика и размеченный набор данных, состоящий из характеристик потоков и меток потоков, которые нужны для обучения и валидации гибридной модели глубокого обучения. При этом наборы данных в БД периодически обновляются.

Основной процесс работы системы может быть представлен следующими этапами.

Этап 1. При поступлении нового потока коммутатор отправляет в контроллер информацию о нем (протокол, порт источника, порт получателя, время прибытия потока, продолжительность потока, интервалы прибытия пакетов в потоке и т. д.).

Этап 2. При получении этой информации на контроллере происходит выделение и первичная обработка данных для передачи их в модуль мониторинга сети.

Этап 3. На основе данных из контроллера модуль мониторинга сети периодически обновляет информацию о потоках в базе данных. В дальнейшем информация передается в модуль статистики потока.

Этап 4. Модуль статистики потока извлекает возможные характеристики и метки потока и сохраняет их в базе данных, а также передает в модуль обработки данных.

Этап 5. Данные обрабатываются для построения обучающего сета и классификации приложений.

Этап 6. ГНС обучается, и в результате модуль глубокого обучения может начать классификацию приложения.

Гибридная сетевая модель глубокого обучения

Далее рассмотрим предлагаемый гибридный метод классификации приложений на основе сети глубокого обучения.

Сеть глубокого обучения

Многоуровневый автокодировщик, исходя из определения, состоит из нескольких архитектур, и представляет собой одну специальную нейронную сеть, состоящую из входного слоя, скрытого слоя и выходного слоя. Типичная архитектура приведена на рисунке 2, где x_i – многомерный вектор входных признаков автокодировщика, y_i – выходные данные скрытого слоя автокодировщика, z_i – выходные данные выходного слоя автокодировщика.

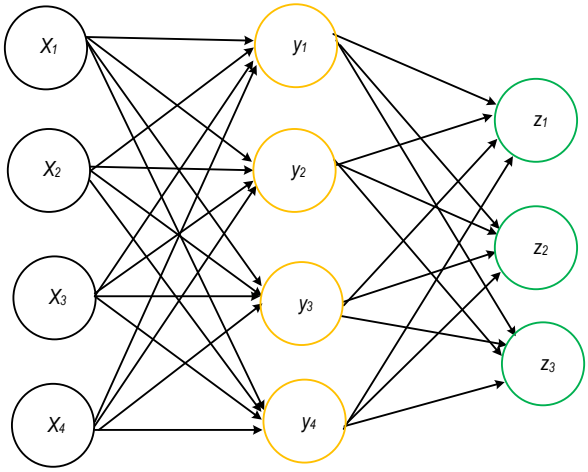


Рис. 2. Модель многоуровневого автокодировщика

Fig. 2 The Stacked Autoencoder Model

В отличие от традиционной модели нейронной сети, выход слоя автокодировщика используется как вход на следующий. В частности, преобразование отображения из входного слоя в скрытый уровень можно рассматривать как процесс кодирования, а отображение данных из скрытого слоя в выходной уровень – это декодирование.

В данном случае x – это M-мерный входной вектор признаков автокодировщика. При подаче необработанных данных на входной слой, их закодированное представление поступает на скрытый слой и может быть представлено по формуле:

$$y(x) = h(W_1 x + b_1). \quad (2)$$

Аналогичным образом можно реконструировать (декодировать) данные для выходного слоя, воспользовавшись выражением:

$$z(x) = f(W_2 y(x) + b_2), \quad (3)$$

где W_1, W_2 – матрицы весов связей между входным и скрытым слоем и между скрытым и выходным слоем в автокодировщике, соответственно; b_1, b_2 – смещение скрытого и выходного слоев в автокодировщике; $h(.)$ – функция активации скрытого слоя в автокодировщике; $f(.)$ – функция активации выходного слоя в автокодировщике.

Целевая функция автокодировщика, необходимая для достижения минимизации ошибки между потоком входных данных и выходных результатов автокодировщика, представлена формулой:

$$L(x, z) = \frac{1}{2N} \sum_{i=1}^{N} \|x_i - z(x_i)\|^2, \quad (4)$$

где N – общее количество обучающих выборок.

Выходные данные последнего скрытого слоя в многоуровневом автокодировщике представляют собой форму входных данных для выходного слоя, поэтому их нельзя использовать для исполнения функции классификации. В данной статье предложено объединить классификатор softmax с многоуровневым автокодировщиком для классификации приложений. Общая модель ГНС для классификации приложений состоит из одного входного слоя, нескольких скрытых слоев и классификатора softmax, как показано на рисунке 3.

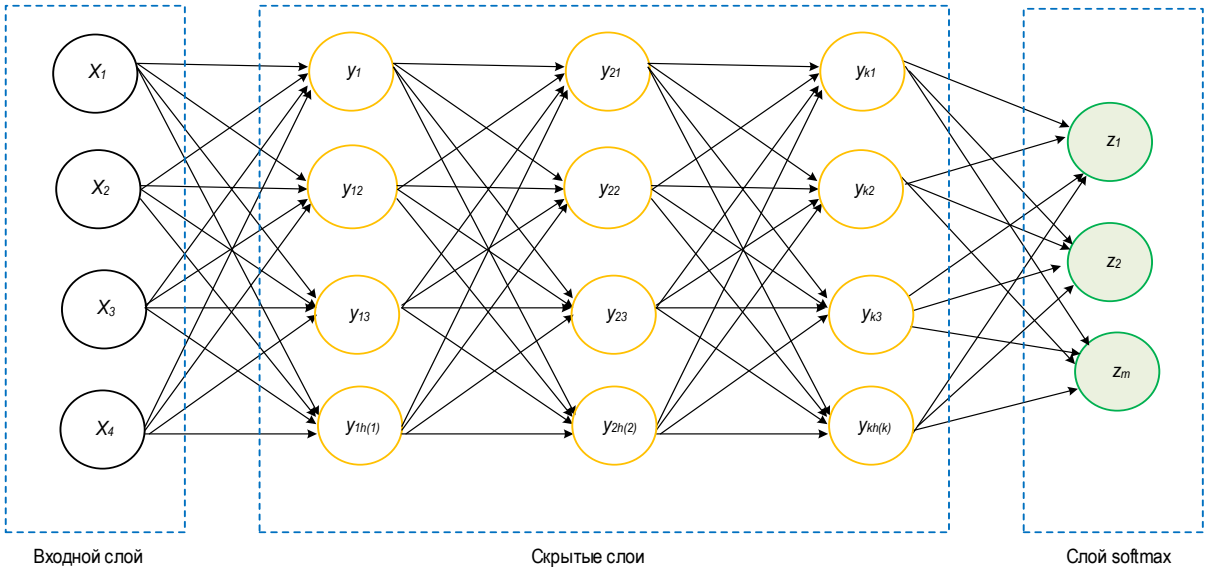


Рис. 3. Модель сети глубокого обучения

Fig. 3. Deep Learning Network Model

В целях обработки неразмеченного набора данных предлагается использование «жадного» алгоритма послойного предобучения [33]. В рамках работы «жадного» алгоритма принимаются локально оптимальные решения на каждом слое, при этом последовательно происходит обучение одного за другим скрытого слоя. Через обучение мы получаем веса и смещения следующего скрытого слоя.

Для предлагаемой модели гибридной DL-сети в процессе обучения используются как индексированный, так и неразмеченный набор данных. Обучение многоуровнего автокодировщика использует немаркированный набор данных. Проиндексированный набор данных необходим для процесса обучения слоя регрессии softmax и процесса тонкой настройки модели гибридной глубокой сети. Так, например: предположим, что помеченный набор данных определяется как X_L = = (X_{L1}, X_{L2}, …, X_{LM}) с M метками Y_L = (Y_{L1}, Y_{L2}, …, Y_{LM}), а немаркированный – как X_U = (X_{U1}, X_{U2}, …, X_{UN}) без N меток, где L –размеченные данные, а U – неразмеченные.

В качестве функции активации для $h(.)$ и $f(.)$ будем использовать сигмоидальную функцию:

$$h(x) = f(x) = S(x) = \frac{1}{1 + e^{-x}}. \quad (5)$$

Повышение эффективности обобщения многоуровневого автокодировщика может быть получено благодаря введению в функцию потерь ограничений по разреженности данных. Таким образом, функция потерь (*LS, от англ.* Lose Function) многоуровневого автокодировщика может быть представлена следующим образом:

$$LS(\Theta_1) = LA\big(x, z(x)\big) + \alpha \sum_{j=1}^{H} KL(\rho \,\|\widehat{\rho_j}), \quad (6)$$

$$LS(\Theta_1) = LA\big(x, z(x)\big) + \alpha \sum_{j=1}^{H} KL(\rho \,\|\widehat{\rho_j}), \quad (7)$$

$$KL\big(\rho\|\widehat{\rho_j}\big) = \rho \log\frac{\rho}{\widehat{\rho_j}} + (1-\rho)\log\frac{1-\rho}{1-\widehat{\rho_j}}, \quad (8)$$

$$\widehat{\rho_j} = \frac{1}{N}\sum_{i=1}^{N} Y_{Uj}(X_{Uj}), \quad (9)$$

где ρ – параметр разреженности многоуровневого автокодировщика; ρ_j – среднее значение активации j-го скрытого узла; H – количество скрытых узлов; $KL(.)$ – функция расхождения Кульбака – Лейблера, которая достигает минимального значения равного 0 при $\rho = \widehat{\rho_j}$, а также при $\theta_1 = (W_1, b_1)$, если W_1 и b_1 – вес и вектор смещения многоуровневого автокодировщика, соответственно; $LA(.)$ – ошибка реконструкции по всем обучающим выборкам; α – коэффициент для регулировки штрафа за разреженность.

Окончательный результат многоуровневого автокодировщика подается в качестве входных данных слоя softmax для обучения классификатора. Таким образом, регрессионная модель softmax обучается с использованием алгоритма контролируемого обучения с помеченным набором образцов.

Предполагается, что все размеченные образцы можно разделить на K классов; другими словами, для каждой выборки X_{Li} соответствующая выходная метка Y_{Li} может принимать K различных значений.

Значение вероятности $P(Z_i = j|X_{Li})$, где X_{Li} относится к j-*му* классу, может быть рассчитано по формуле (10). При этом $\Theta_2 = (W_2, b_2)$, где W_2 и b_2 – это вес и вектор смещения слоя регрессии softmax, соответственно.

Функция потерь классификатора softmax определяется согласно выражению (11), где M – общее количество наборов данных; $1(Z_i = j)$ – дискретная функция, которая принимает значение 1 при выполнении определенного условия, в противном случае принимает значение, равное 0; β – штрафной коэффициент; n – количество узлов входного слоя; θ_{ij} – вектор веса и смещения слоя softmax.

В соответствии с данной схемой можно определить целевую функцию гибридной модели на этапе тонкой настройки, как показано ниже, а третье слагаемое представляет собой штрафную функцию модели многоуровневого автокодировщика, согласно выражению (12) при $\theta = \theta_1, \theta_2$, где H – общее количество скрытых слоев многоуровневого автокодировщика; θ_h – вектор веса и смещения многоуровневой модели автокодировщика.

$$h_{\Theta_2}\big(X_{L_i}\big) = \begin{bmatrix} P\big(Z_1 = 1\big|X_{L_i};\Theta_2\big) \\ P\big(Z_2 = 2\big|X_{L_i};\Theta_2\big) \\ \vdots \\ P\big(Z_K = K\big|X_{L_i};\Theta_2\big) \end{bmatrix} = \frac{1}{\sum_{j=1}^{K} e^{\Theta_j^T X_{L_i}}} \begin{bmatrix} e^{\Theta_1^T X_{L_i}} \\ e^{\Theta_2^T X_{L_i}} \\ \vdots \\ e^{\Theta_K^T X_{L_i}} \end{bmatrix}. \quad (10)$$

$$LSM(\Theta_2) = -\frac{1}{M}\left\{\sum_{i=1}^{M}\sum_{j=1}^{K} 1(Z_i = j)\log\frac{e^{\Theta_j^T X_{L_i}}}{\sum_{S=1}^{K} e^{\Theta_S^T X_{L_i}}}\right\} + \frac{\beta}{2}\sum_{i=1}^{K}\sum_{j=1}^{n}\Theta_{ij}^2\, 11\,. \quad (11)$$

$$LSS(\theta) = -\frac{1}{M}\left\{\sum_{i=1}^{M}\sum_{j=1}^{K} 1(Z_i = j)\log\frac{e^{\Theta_j^T X_{L_i}}}{\sum_{S=1}^{K} e^{\Theta_S^T X_{L_i}}}\right\} + \frac{\gamma}{2}\sum_{i=1}^{K}\sum_{j=1}^{n}\Theta_{ij}^2 + \frac{\gamma}{2}\sum_{h=1}^{H}\Theta_h^2\,. \quad (12)$$

Процесс обучения

Процесс обучения модели сети глубокого обучения состоит из 3-х этапов: предварительного обучения многоуровневого автокодировщика, обучения слоя softmax и тонкой настройки общей ГНС. На начальном этапе «жадный» алгоритм послойного предобучения используется для обучения весов многоуровневого автокодировщика методом без учителя. При обучении слоя регрессии softmax уже используется алгоритм с учителем. Во время тонкой настройки ГНС применяется алгоритм обратного распространения ошибки для точной настройки сети глубокого обучения, используя метод по размеченным данным. Таким образом, процесс обучения можно представить следующими этапами.

Этап 1. Немаркированные наборы данных X_U применяются для обучения первого скрытого слоя модели многоуровневого автокодировщика на основе метода без учителя путем минимизации целевой функции $LS(\theta_1)$.

Этап 2. Выходные данные первого скрытого слоя используются в качестве входных данных второго скрытого слоя для обучения второго автокодировщика в рамках той же целевой функцией $LS(\theta_1)$.

Этап 3. Выходные данные i-го скрытого слоя используются как входные данные $(i + 1)$-го скрытого слоя для обучения $(i + 1)$-го автокодировщика. Данная операция повторяется до тех пор, пока последний скрытый слой не будет достаточно обучен.

Этап 4. Выходные данные последнего скрытого слоя подаются в качестве входных данных слоя softmax для его обучения с помощью размеченного набора выборок данных X_L, используя алгоритм обучения с учителем.

Этап 5. Модель ГНС с полученными весами и смещением из вышеупомянутого процесса обучения активируется, а затем размеченный набор образцов X_L используется для точной настройки общей ГНС. Это минимизирует целевую функцию $LSS(\theta)$ с помощью алгоритма обратного распространения ошибки.

Алгоритм обучения для всей глубокой нейронной сети проиллюстрирован на рисунке 4. На первом шаге параметры многоуровневого автокодировщика включают веса W, смещение b, количество скрытых слоев M и комбинацию скрытых узлов в каждом скрытом слое. На десятом шаге параметры в слое softmax состоят из весов и смещения первичных для этого слоя. После этого на шаге 12 происходит использование параметров, полученных в процессе предварительной подготовки, для обновления весов гибридной ГНС, а алгоритм обратного распространения ошибки затем точнее настраивает веса во всей модели.

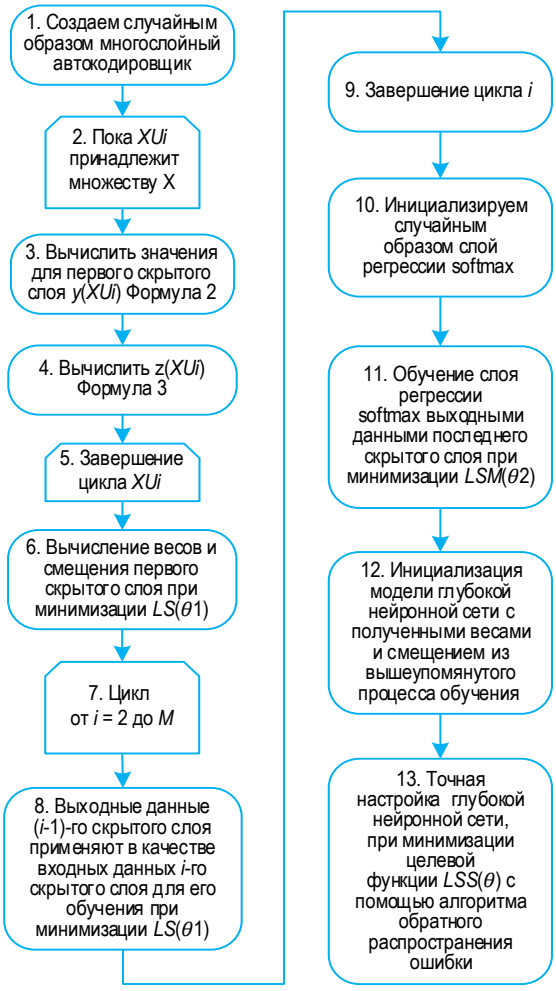


Рис. 4. Алгоритм обучения ГНС

Fig. 4. Deep Neural Network Learning Algorithm

Экспериментальные исследования

Для оценки производительности предлагаемого метода классификации приложений был проведен эксперимент по моделированию, который реализован в программной платформе Weka и MATLAB, работающей на персональном компьютере Intel (R) Core™ / 2,93 ГГц / 4 ГБ под управлением операционной системы Windows.

При моделировании каждый эксперимент повторялся 100 раз и бралось среднее значение в качестве окончательного результата. Коэффициенты ρ, α, β и γ инициализируются равными 0.1, 3, 0.001 и 0.001, соответственно. Кроме того, мы устанавливаем число итерации обучения равным 3000. Перед предварительной тренировкой случайным образом генерируются начальные векторы веса и смещения многоуровневого автокодировщика и слоя регрессии softmax.

Оценка характеристик предложенного метода классификации, с предлагаемым методом и моделью глубокого обучения будет производиться в сравнении с традиционной моделью классификатора на базе SVM.

Набор данных

Для тестирования метода был взят набор данных, полученный в компьютерной лаборатории Кембриджского университета и широко используемый во многих исследовательских работах по классификации трафика. [4, 8, 10, 25]. Набор данных состоит из отдельных файлов, собранных в разное время дня, и каждый набор – только из потоков TCP-трафика, для которых выделено 249 характеристик: в частности, продолжительность потока, метка класса приложения и др. Все сетевые потоки подразделяются на 10 классов, что приведено ниже в нотации «Условный номер: тип трафика / приложения (протоколы)»: 1) Передача данных / ftp; 2) Базы данных / postgress, sqlnet oracle, ingress; 3) Служебный трафик / Ssh, klogin, rlogin, telnet; 4) Почта / imap, pop2/3, smtp; 5) Сервисы / X11, dns, ident, ldap, ntp; 6) Трафик Интернет / www; 7) P2P / KaZaA, BitTorrent, GnuTella; 8) Злонамеренный трафик / Сигнатуры сетевых атак; 9) Игры / Microsoft Direct Play; 10) Мультимедиа / Windows Media Player, Real.

Исходные наборы данных можно скачать с сайта отдельного проекта компьютерной лаборатории Кембриджского университета [34].

Оценка работы модели

Чтобы оценить эффективность и производительность предложенного метода классификации, рассматривались такие показатели, как вероятность верного срабатывания, точность и полнота.

Вероятность верной классификации – это отношение общего количества приложений, правильно типизированные классификатором, к общему количеству образцов приложений. Используется для оценки точности классификатора приложений на всем наборе данных.

Точность классификации – это доля приложений, которые правильно отнесено к данному классу приложений.

Полнота классификации – это доля приложений, принадлежащих классу *i*, которые правильно отнесены к конкретному классу *i*.

Поскольку гибридная глубокая нейронная сеть с 5 скрытыми слоями и 10 скрытыми узлами имеет самую высокую вероятность верной классификации, поэтому в качестве основной DL-модели будет использоваться именно она (таблица 1).

На рисунке 5a (слева) приведено сравнение вероятностей верной классификации приложений для DL-моделей и на базе SVM для 10 различных наборов данных. Из сравнения видно, что первая имеет более высокую точность классификации. Однако стоит отметить, что DL-модель состоит из бóльшего количества скрытых слоев и использует алгоритм обучения более высокого уровня, поэтому способность к обучению классификации получается более эффективной. Кроме того, поскольку распределение трафика в каждом наборе данных отличается, общая точность классификации для разных наборов данных также незначительно разнится.

ТАБЛИЦА 1 Значения вероятности верной классификации в рассматриваемой модели

TABLE 1. The Performance of Deep Learning Network

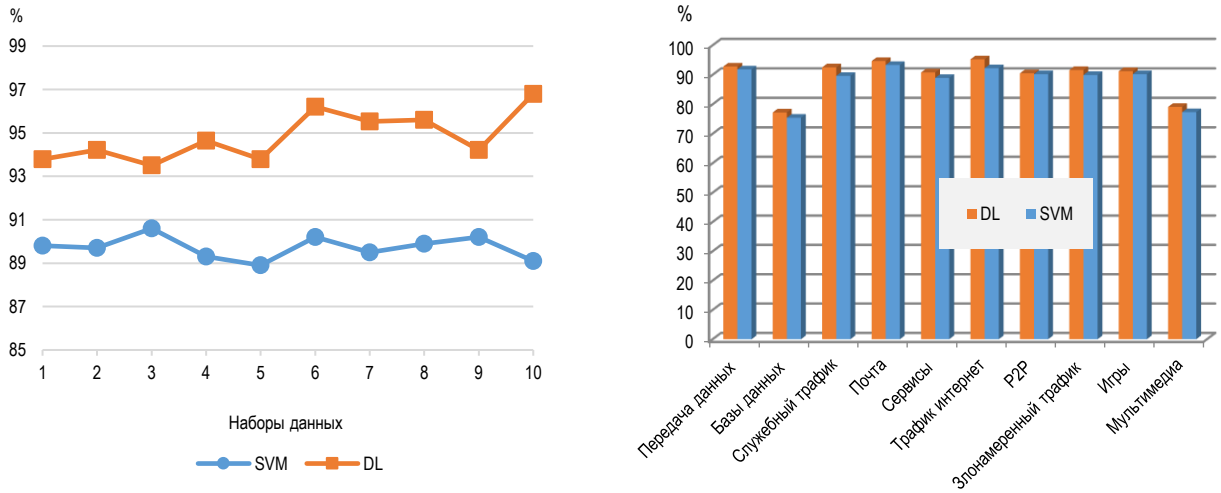
№ п/п	Число скрытых слоев	Число скрытых узлов	Вероятность верного срабатывания, %
1	3	5	82.44
2	3	10	85.05
3	3	15	85.33
4	4	5	85.63
5	4	10	88.75
6	4	15	88.15
7	5	5	90.9
8	5	10	91.55
9	5	15	90.2
10	6	5	90.77
11	6	10	88.04
12	6	15	86.31

На рисунке 5a (справа) показана вероятность верной классификации приложений для моделей DL и на базе SVM (список классов приложений приведен выше). Можно отметить, что точность классификации для трафика Интернет, почты, передачи данных, служебного трафика и сервисов относительно высока, независимо от того, какая модель используется. В связи с тем, что количество таких приложений в выборках составляет большинство в каждом наборе данных, у моделей больше данных для их идентификации.

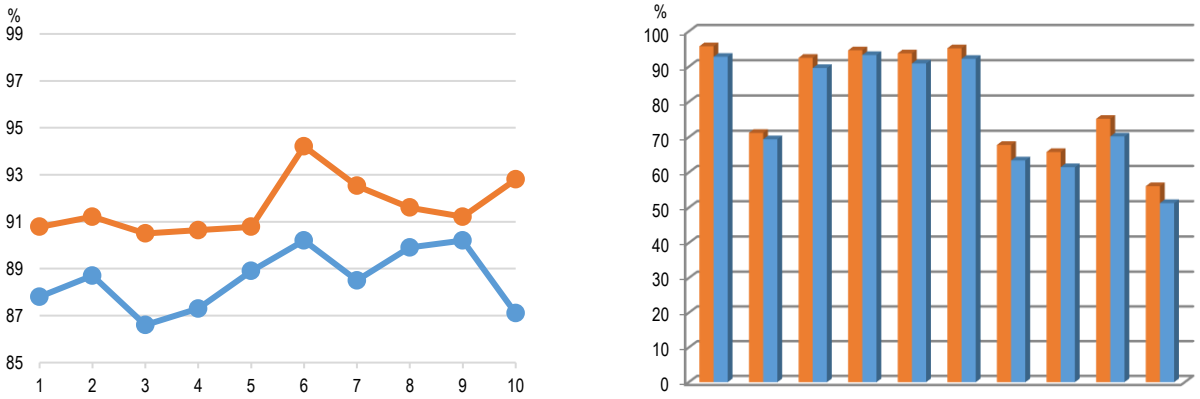
Кроме того, гибридная модель DL-сети, состоящая из многоуровневого автокодировщика и слоя регрессии softmax, более эффективна с точки зрения обучения и возможностей классификации. При этом стоит отметить, что отдельные классы трафика, такие как мультимедиа и базы данных, имеют невысокую вероятность верной классификации, что объясняется небольшим числом пакетов этих приложений в используемых наборах данных. Аналогичные графики для определения точности классификации (рисунок 5b) и полноты классификации (рисунок 5c) приведены ниже.

На рисунке 5c (справа) можно констатировать, что полнота классификации DL-модели выше, чем на базе SVM для различных сетевых приложений. Возможно, это связано с бóльшим числом скрытых слоев, которые обеспечивают дополнительные возможности к обобщению и принятию решения, а также эффективности извлечения признаков и абстракции приложений в трафике.

a) Вероятность верной классификации
Accuracy Comparison



b) Точность классификации
Precision Comparison



c) Полнота классификации
Recall Comparison

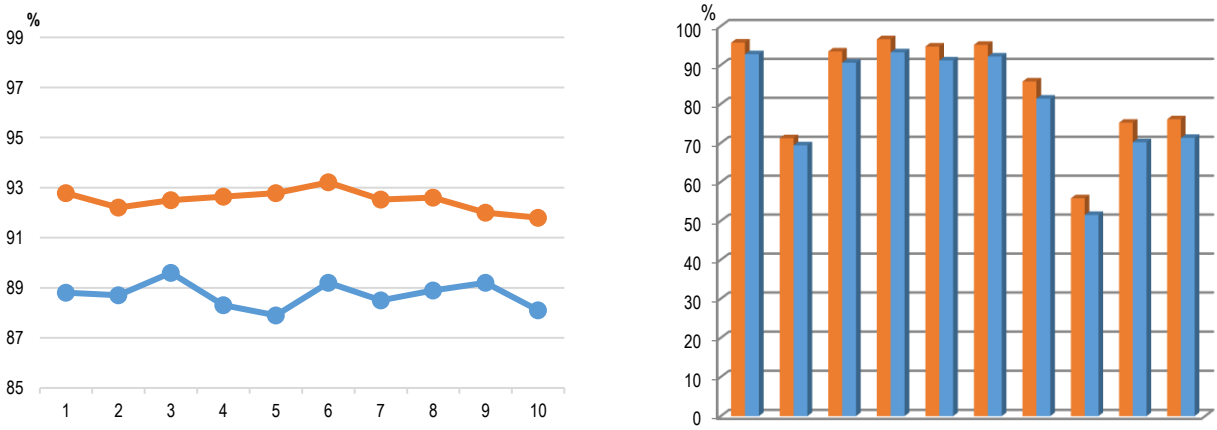


Рис. 5. Сравнение классификаций характеристик для моделей глубокого обучения и на базе SVM (для различных наборов – слева, в рамках каждого класса – справа)

Fig 5. Comparison between Deep Learning and Support Vector Machine: (left – for different data sets, right – for different applications)

Заключение

Применение архитектуры SDN позволило разработать эффективное применение новой системы классификации сетевых приложений на основе гибридной сети глубокого обучения, состоящей из многоуровневого автокодировщика и слоя регрессии softmax. Благодаря логическому централизованному управлению в SDN и мощным вычисли-

тельным возможностям потоки сетевых приложений могут собираться и обрабатываться в контроллере. В дальнейшем простой многоуровневый автокодировщик используется для получения признаков потока, а также выделения признаков более высокого уровня, а уровень регрессии softmax – в качестве классификатора для итоговой идентификации сетевых приложений.

Отдельно стоит отметить применение немаркированных данных для обучения многоуровневого автокодировщика и набора размеченных данных для обучения слоя softmax, а также эффективное применение метода обратного распространения ошибки для тонкой донастройки всей гибридной ГНС.

Несмотря на то, что в статье предложена структура классификации приложений, основанная на архитектуре SDN и технологии DL, основной акцент сделан на подтверждение эффективности и оценку предложенного метода классификации приложений на основе глубокого обучения в соответствии с имеющимся набором данных. Результаты экспериментов указывают на более высокие комплексные характеристики точности классификации по сравнению с моделью на базе SVM.

С развитием гетерогенных сетей, при взрывном росте сетевого трафика, сетевые приложения становятся все более диверсифицированными, и, хотя предлагаемый метод обеспечивает классификацию приложений, остаются нерешенными отдельные прикладные вопросы. Один из них связан с созданием предварительных размеченных наборов сетевого трафика, получение которых в реальной сети, с учетом большого числа разнообразных приложений, достаточно сложно. Поэтому в дальнейшем рассматривается возможность использования алгоритма глубокого обучения без учителя для классификации приложений. Кроме того, необходимо снизить временные затраты на обучение ГНС, которые на данный момент достаточно велики, по сравнению с сетевой задержкой.

Список источников

1. Елагин В.С. Динамическое управление нагрузкой в программно-конфигурируемых сетях // Труды учебных заведений связи. 2017. Т. 3. № 3. С. 60–67.

2. Елагин В.С., Дмитриева Ю.С. Моделирование сетевого ресурса в программно-конфигурируемых сетях // Вестник связи. 2020. № 6. С. 35–40.

3. Zhang J., Chen X., Xiang Y., Zhou W., Wu J. Robust Network Traffic Classification // IEEE /ACM Transactions on Networking. 2015. Vol. 23. Iss. 4. PP. 1257–1270. DOI:10.1109/TNET.2014.2320577

4. Kim H., Claffy K.C., Fomenkov M., Barman D., Faloutsos M., Lee K. Internet traffic classification demystified: myths, caveats, and the best practices // Proceedings of the Conference on emerging Networking EXperiments and Technologies (Madrid, Spain, 9–12 December 2008). New York: Association for Computing Machinery, 2008. DOI:10.1145/1544012.1544023

5. Auld T., Moore A.W., Gull S.F. Bayesian Neural Networks for Internet Traffic Classification // IEEE Transactions Neural Networ. 2007. Vol. 18. Iss. 1. PP. 223–239. DOI:10.1109/TNN.2006.883010

6. Nguyen T.T.T., Armitage G. A survey of techniques for internet traffic classification using machine learning // IEEE Communication Survive Tutorials. 2008. Vol. 10. Iss. 4. PP. 56–76. DOI:10.1109/SURV.2008.080406

7. Valenti S., Rossi D., Dainotti A., Pescapè A., Finamore A., Mellia M. Reviewing Traffic Classification // Biersack E., Callegari C., Matijasevic M. (eds) Data Traffic Monitoring and Analysis. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2013. Vol. 7754. PP. 123–147. DOI:10.1007/978-3-642-36784-7_6

8. Zhang J., Chen C., Xiang Y., Zhou W., Xiang Y. Internet Traffic Classification by Aggregating Correlated Naive Bayes Predictions // IEEE Transactions on Information Forensics and Security. 2013. Vol. 8. Iss. 1. PP. 5–15. DOI:10.1109/TIFS.2012.2223675

9. Grimaudo L., Mellia M., Baralis E., Keralapura R. SeLeCT: Self-Learning Classifier for Internet Traffic // IEEE Transactions Network Service Management. 2014. Vol. 11. Iss. 2. PP. 144–157. DOI:10.1109/TNSM.2014.011714.130505

10. Cao J., Fang Z., Qu G., Sun H., Zhang D. An accurate traffic classification model based on support vector machines // International Journal of Network Management. 2017. Vol. 27. Iss. 1. P. e1962. DOI:10.1002/nem.1962

11. Pasca S.T.V., Prasad S.S., Kataoka K. AMPF: Application-aware Multipath Packet Forwarding using Machine Learning and SDN // arXiv:1606.05743. 2016. DOI:10.48550/arXiv.1606.05743

12. Amaral P., Dinis J., Pinto P., Bernardo L., Tavares J., Mamede H.S. Machine Learning in Software Defined Networks: Data Collection and Traffic Classification // Proceedings of the 24th International Conference on Network Protocols (ICNP, Singapore, 08–11 November 2016). IEEE, 2016. DOI:10.1109/ICNP.2016.7785327

13. Wang P., Lin S.C., Luo M. A Framework for QoS-aware Traffic Classification Using Semi-supervised Machine Learning in SDNs // Proceedings of the International Conference on Services Computing (SCC, San Francisco, USA, 27 June – 02 July 2016). IEEE, 2016. DOI:10.1109/SCC.2016.133

14. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. 2015. Vol. 521. Iss. 7553. PP. 436–444. DOI:10.1038/nature14539

15. Chen X.W., Lin X. Big Data Deep Learning: Challenges and Perspectives // IEEE Access. 2014. Vol. 2. PP. 514–525. DOI:10.1109/ACCESS.2014.2325029

16. Kreutz D., Ramos F.M.V., Verissimo P.E., Rothenberg C.E., Azodolmolky S., Uhlig S. Software-Defined Networking: a Comprehensive Survey // Proceedings of the IEEE. 2015. Vol. 103. Iss. 1. PP. 14–76. DOI:10.1109/JPROC.2014.2371999

17. Bu C., Wang X., Cheng H., Huang M., Li K., Das S. Enabling Adaptive Routing Service Customization via the Integration of SDN and NFV // Journal of Network Computing Applications. 2017. Vol. 93. PP. 123–136. DOI:10.1016/j.jnca.2017.05.010

18. Yi B., Wang X., Huang M. Design and evaluation of schemes for provisioning service function chainwith function scalability // Journal of Network Computing Applications. 2017. Vol. 93. PP. 197–214. DOI:10.1016/j.jnca.2017.05.013
19. Lv J., Wang X., Huang M., Shi J., Li K., Li J. RISC: ICN routing mechanism incorporating SDN and community division // Computing Network. 2017. Vol. 123. PP. 88–103. DOI:10.1016/j.comnet.2017.05.010
20. He Q., Wang X., Huang M. OpenFlow-based low-overhead and high-accuracy SDN measurement framework // Transactions on Emerging Telecommunications Technologies. 2018. Vol. 29. Iss. 2. P. e3263. DOI:10.1002/ett.3263
21. Yi B., Wang X., Li K., Das S.K., Huang M. A comprehensive survey of Network Function Virtualization // Computing Network. 2018. Vol. 133. PP. 212–262. DOI:10.1016/j.comnet.2018.01.021
22. Shu Z., Wan J., Lin J., Wang S., Li D., Rho S., et al. Traffic engineering in software-defined networking: Measurement and management // IEEE Access. 2016. Vol. 4. PP. 3246–3256. DOI:10.1109/ACCESS.2016.2582748
23. Cui L., Yu F.R., Yan Q. When big data meets software-defined networking: SDN for big data and big data for SDN // IEEE Network. 2016. Vol. 30. Iss. 1. PP. 58–65. DOI:10.1109/MNET.2016.7389832
24. Zhang L., Huang H., Jing X. A modified cyclostationary spectrum sensing based on softmax regression model // Proceedings of the 16th International Symposium on Communications and Information Technologies (ISCIT, Qingdao, China, 26–28 September 2016). IEEE, 2016. DOI:10.1109/ISCIT.2016.7751707
25. Zhang H., Lu G., Qassrawi M.T., Zhang Y., Yu X. Feature selection for optimizing traffic classification // Computing Communicdtion. 2012. Vol. 35. Iss. 12. PP. 1457–1471. DOI:10.1016/j.comcom.2012.04.012
26. da Silva A.S., Machado C.C., Bisol R.V., Granville L.Z., Schaeffer A. Identification and Selection of Flow Features for Accurate Traffic Classification in SDN // Proceedings of the 14th International Symposium on Network Computing and Applications (NCA, Cambridge, USA, 28–30 September 2015). IEEE, 2015. DOI:10.1109/NCA.2015.12
27. Schmidhuber J. Deep learning in neural networks: an overview // Neural Network. 2015. Vol. 61. PP. 85–117. DOI:10.1016/j.neunet.2014.09.003
28. Salama M.A., Eid H.F., Ramadan R.A., Darwish A., Hassanien E. Hybrid Intelligent Intrusion Detection Scheme // Gaspar-Cunha A., Takahashi R., Schaefer G., Costa L. (eds) Soft Computing in Industrial Applications. Advances in Intelligent and Soft Computing. Berlin, Heidelberg: Springer, 2011. Vol. 96. PP. 293–303. DOI:10.1007/978-3-642-20505-7_26
29. Fiore U., Palmieri F., Castiglione A., De Santis A. Network anomaly detection with the restricted Boltzmann machine // Neurocomputing. 2013. Vol. 122. PP. 13–23. DOI:10.1016/j.neucom.2012.11.050
30. Lv Y., Duan Y., Kang W., Li Z., Wang F.Y. Traffic Flow Prediction with Big Data: a Deep Learning Approach // IEEE Transactions Intelligent Transport System. 2015. Vol. 16. Iss. 2. PP. 865–873. DOI:10.1109/TITS.2014.2345663
31. Yang H.F., Dillon T.S., Chen Y.P. Optimized Structure of the Traffic Flow Forecasting Model with a Deep Learning Approach // IEEE Transactions on Neural Networks and Learning Systems. 2017. Vol. 28. Iss. 10. PP. 2371–2381. DOI:10.1109/TNNLS.2016.2574840
32. Huang W., Song G., Hong H., Xie K. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks with Multitask Learning // IEEE Transactions Intelligent Transport System. 2014. Vol. 15. Iss. 5. PP. 2191–2201. DOI:10.1109/TITS.2014.2311123
33. Bengio Y., Lamblin P., Popovici D., Larochelle H. Greedy Layer-Wise Training of Deep Networks // Proceedings of the Conference on Advances in Neural Information Processing Systems 19 (2006). MIT Press, 2007. PP. 153–160.
34. BRASIL. Characterizing Network-based Applications. Data sets // University of Cambridge Computer Laboratory. URL: https://www.cl.cam.ac.uk/research/srg/netos/projects/brasil/data/index.html (дата обращения 15.06.2023)

References

1. Elagin V. Dynamic Load Balancing in Software-Defined Network. Proceedings of Telecommun. Univ. 2017;3(3):60–67.
2. Elagin V.S., Dmitrieva Yu.S. The Modeling of Network Resources in Software-Defined Networks. *Vestnik Communications.* 2020;6:35–40.
3. Zhang J., Chen X., Xiang Y., Zhou W., Wu J. Robust Network Traffic Classification. *IEEE /ACM Transactions on Networking*. 2015;23(4):1257–1270. DOI:10.1109/TNET.2014.2320577
4. Kim H., Claffy K.C., Fomenkov M., Barman D., Faloutsos M., Lee K. Internet traffic classification demystified: myths, caveats, and the best practices. *Proceedings of the Conference on emerging Networking EXperiments and Technologies, 9–12 December 2008, Madrid, Spain*. New York: Association for Computing Machinery; 2008. DOI:10.1145/1544012.1544023
5. Auld T., Moore A.W., Gull S.F. Bayesian Neural Networks for Internet Traffic Classification. *IEEE Transactions Neural Networ*. 2007;18(1):223–239. DOI:10.1109/TNN.2006.883010
6. Nguyen T.T.T., Armitage G. A survey of techniques for internet traffic classification using machine learning. *IEEE Communication Survive Tutorials*. 2008;10(4):56–76. DOI:10.1109/SURV.2008.080406
7. Valenti S., Rossi D., Dainotti A., Pescapè A., Finamore A., Mellia M. Reviewing Traffic Classification. *In: Biersack E., Callegari C., Matijasevic M. (eds) Data Traffic Monitoring and Analysis. Lecture Notes in Computer Science, vol.7754*. Berlin, Germany: Springer; 2013. p.123–147. DOI:10.1007/978-3-642-36784-7_6
8. Zhang J., Chen C., Xiang Y., Zhou W., Xiang Y. Internet Traffic Classification by Aggregating Correlated Naive Bayes Predictions. *IEEE Transactions on Information Forensics and Security*. 2013:8(1):5–15. DOI:10.1109/TIFS.2012.2223675
9. Grimaudo L., Mellia M., Baralis E., Keralapura R. SeLeCT: Self-Learning Classifier for Internet Traffic. *IEEE Transactions Network Service Management*. 2014;11(2):144–157. DOI:10.1109/TNSM.2014.011714.130505
10. Cao J., Fang Z., Qu G., Sun H., Zhang D. An accurate traffic classification model based on support vector machines. *International Journal of Network Management*. 2017;27(1):e1962. DOI:10.1002/nem.1962
11. Pasca S.T.V., Prasad S.S., Kataoka K. AMPF: Application-aware Multipath Packet Forwarding using Machine Learning and SDN. *arXiv:1606.05743*. 2016. DOI:10.48550/arXiv.1606.05743

12. Amaral P., Dinis J., Pinto P., Bernardo L., Tavares J., Mamede H.S. Machine Learning in Software Defined Networks: Data Collection and Traffic Classification. *Proceedings of the 24th International Conference on Network Protocols, ICNP, 08–11 November 2016, Singapore*. IEEE; 2016. DOI:10.1109/ICNP.2016.7785327
13. Wang P., Lin S.C., Luo M. A Framework for QoS-aware Traffic Classification Using Semi-supervised Machine Learning in SDNs. *Proceedings of the International Conference on Services Computing, SCC, 27 June – 02 July 2016, San Francisco, USA*. IEEE; 2016. DOI:10.1109/SCC.2016.133
14. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444. DOI:10.1038/nature14539
15. Chen X.W., Lin X. Big Data Deep Learning: Challenges and Perspectives. *IEEE Access*. 2014;2:514–525. DOI:10.1109/ACCESS.2014.2325029
16. Kreutz D., Ramos F.M.V., Verissimo P.E., Rothenberg C.E., Azodolmolky S., Uhlig S. Software-Defined Networking: a Comprehensive Survey. *Proceedings of the IEEE*. 2015;103(1):14–76. DOI:10.1109/JPROC.2014.2371999
17. Bu C., Wang X., Cheng H., Huang M., Li K., Das S. Enabling Adaptive Routing Service Customization via the Integration of SDN and NFV. *Journal of Network Computing Applications*. 2017;93:123–136. DOI:10.1016/j.jnca.2017.05.010
18. Yi B., Wang X., Huang M. Design and evaluation of schemes for provisioning service function chainwith function scalability. *Journal of Network Computing Applications*. 2017;93:197–214. DOI:10.1016/j.jnca.2017.05.013
19. Lv J., Wang X., Huang M., Shi J., Li K., Li J. RISC: ICN routing mechanism incorporating SDN and community division. *Computing Network*. 2017;123:88–103. DOI:10.1016/j.comnet.2017.05.010
20. He Q., Wang X., Huang M. OpenFlow-based low-overhead and high-accuracy SDN measurement framework. *Transactions on Emerging Telecommunications Technologies*. 2018;29(2):e3263. DOI:10.1002/ett.3263
21. Yi B., Wang X., Li K., Das S.K., Huang M. A comprehensive survey of Network Function Virtualization. *Computing Network*. 2018;133:212–262. DOI:10.1016/j.comnet.2018.01.021
22. Shu Z., Wan J., Lin J., Wang S., Li D., Rho S., et al. Traffic engineering in software-defined networking: Measurement and management. *IEEE Access*. 2016;4:3246–3256. DOI:10.1109/ACCESS.2016.2582748
23. Cui L., Yu F.R., Yan Q. When big data meets software-defined networking: SDN for big data and big data for SDN. *IEEE Network.* 2016;30(1):58–65. DOI:10.1109/MNET.2016.7389832
24. Zhang L., Huang H., Jing X. A modified cyclostationary spectrum sensing based on softmax regression model. *Proceedings of the 16th International Symposium on Communications and Information Technologies, ISCIT, 26–28 September 2016, Qingdao, China*. IEEE; 2016. DOI:10.1109/ISCIT.2016.7751707
25. Zhang H., Lu G., Qassrawi M.T., Zhang Y., Yu X. Feature selection for optimizing traffic classification. *Computing Communicdtion*. 2012;35(12):1457–1471. DOI:10.1016/j.comcom.2012.04.012
26. da Silva A.S., Machado C.C., Bisol R.V., Granville L.Z., Schaeffer A. Identification and Selection of Flow Features for Accurate Traffic Classification in SDN. *Proceedings of the 14th International Symposium on Network Computing and Applications, USA, 28–30 September 2015, NCA, Cambridge*. IEEE; 2015. DOI:10.1109/NCA.2015.12
27. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Network*. 2015;61:85–117. DOI:10.1016/j.neunet.2014.09.003
28. Salama M.A., Eid H.F., Ramadan R.A., Darwish A., Hassanien E. Hybrid Intelligent Intrusion Detection Scheme. *In: Gaspar-Cunha A., Takahashi R., Schaefer G., Costa L. (eds) Soft Computing in Industrial Applications. Advances in Intelligent and Soft Computing, vol.96*. Berlin, Heidelberg: Springer; 2011. p.293–303. DOI:10.1007/978-3-642-20505-7_26
29. Fiore U., Palmieri F., Castiglione A., De Santis A. Network anomaly detection with the restricted Boltzmann machine. *Neurocomputing*. 2013;122:13–23. DOI:10.1016/j.neucom.2012.11.050
30. Lv Y., Duan Y., Kang W., Li Z., Wang F.Y. Traffic Flow Prediction with Big Data: a Deep Learning Approach. *IEEE Transactions Intelligent Transport System*. 2015;16(2):865–873. DOI:10.1109/TITS.2014.2345663
31. Yang H.F., Dillon T.S., Chen Y.P. Optimized Structure of the Traffic Flow Forecasting Model with a Deep Learning Approach. *IEEE Transactions on Neural Networks and Learning Systems*. 2017;28(10):2371–2381. DOI:10.1109/TNNLS.2016.2574840
32. Huang W., Song G., Hong H., Xie K. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks with Multitask Learning. *IEEE Transactions Intelligent Transport System*. 2014;15(5):2191–2201. DOI:10.1109/TITS.2014.2311123
33. Bengio Y., Lamblin P., Popovici D., Larochelle H. Greedy Layer-Wise Training of Deep Networks. *Proceedings of the Conference on Advances in Neural Information Processing Systems 19, 2006*. MIT Press; 2007. p.153–160.
34. *University of Cambridge Computer Laboratory*. BRASIL. Characterizing Network-based Applications. Data sets. URL: https://www.cl.cam.ac.uk/research/srg/netos/projects/brasil/data/index.html [Accessed 15.06.2023]

Статья поступила в редакцию 26.07.2023; одобрена после рецензирования 21.08.2023; принята к публикации 28.08.2023.
The article was submitted 26.07.2023; approved after reviewing 21.08.2023; accepted for publication 28.08.2023.

Информация об авторе:

ЕЛАГИН Василий Сергеевич
кандидат технических наук, доцент, доцент кафедры Инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0003-4213-953X

Научная статья
УДК 004.42
DOI:10.31854/1813-324X-2023-9-5-79-90

Методология реверс-инжиниринга машинного кода. Часть 1. Подготовка объекта исследования

Константин Евгеньевич Израилов, konstantin.izrailov@mail.ru

Санкт-Петербургский Федеральный исследовательский центр Российской академии наук,
Санкт-Петербург, 199178, Российская Федерация

Аннотация: *Изложены результаты создания единой методологии проведения реверс-инжиниринга машинного кода устройств. В первой части цикла статей проводится обзор научных публикаций данной предметной области. В условиях отсутствия удовлетворительных решений предлагается авторская методология процесса, состоящая из 4-х следующих этапов: подготовительные мероприятия, статическое исследование, динамическое исследование и документирование. Приводится детальное описание шагов 1-го этапа, а также примеры их применения на практике с использованием типовых программных средств. Схема предлагаемой методологии представлена в графическом виде, а приведенные шаги имеют формальную запись. В следующих частях цикла статей будут описаны шаги остальных этапов и их систематизация в табличном виде с указанием входных и выходных объектов, а также формы выполнения шагов.*

Ключевые слова: *реверс-инжиниринг, обратная разработка, программная инженерия, информационная безопасность, уязвимости, методология, IDA Pro*

Ссылка для цитирования: Израилов К.Е. Методология реверс-инжиниринга машинного кода. Часть 1. Подготовка объекта исследования // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 79–90. DOI:10.31854/1813-324X-2023-9-5-79-90

Methodology for Machine Code Reverse Engineering. Part 1. Preparation of the Research Object

Konstantin Izrailov, konstantin.izrailov@mail.ru

Saint-Petersburg Federal Research Center of the Russian Academy of Sciences,
St. Petersburg, 199178, Russian Federation

Abstract: *The results of creating a unified methodology for reverse engineering the devices machine code are presented. The first part of the series of articles reviews scientific publications in this subject area. In the absence of satisfactory solutions, the author's process methodology is proposed, consisting of the following 4 stages: preparatory activities, static research, dynamic research and documentation. A detailed description of the steps of the first stage is provided, as well as examples of their application in practice using standard software. The scheme of the proposed methodology is presented in graphical form, and the steps given are formally written. The next part of the series of articles will describe the steps of the remaining stages and their systematization in tabular form, indicating the input and output objects, as well as the form of steps execution.*

Keywords: *reverse engineering, software engineering, information security, vulnerabilities, methodology, IDA Pro*

For citation: Izrailov K. Methodology for Machine Code Reverse Engineering. Part 1. Preparation of the Research Object. *Proceedings of Telecommun. Univ.* 2023;9(5):79–90. DOI:10.31854/1813-324X-2023-9-5-79-90

1. Введение

Задача исследования машинного кода (далее – МК) программного обеспечения считается актуальной уже многие десятки лет. Основная цель такого исследования ставится как получение информации о функционале кода, отличие которого от заявленного может означать наличие в нем уязвимостей [1]. Также в ряде случаев необходима доработка программы или ее полная замена на аналог, а в случае отсутствия исходных кодов их восстановление из МК (а также алгоритмов и архитектуры) является практически единственным решением [2]. Сам процесс носит название «реверс-инжиниринга» (далее – РИ), заимствованное из английского «reverse engineering», что переводится, как интуитивно понятный термин – «обратная разработка».

Исходя из указанных потребностей, в РИ создано определенное количество методов и программных средств (например, IDA Pro, Ghidra, Radare2 и др.), позволяющих проводить исследование и восстанавливать необходимую информацию для отдельно взятых программ; при этом участие эксперта в процессе велико, поскольку какая-либо полностью специфицированная и автоматическая процедура отсутствует. Успешность же РИ оказывается полностью зависящей от знаний эксперта, его опыта, наличия под рукой средств, а также самого исследуемого экземпляра МК; при этом, если специфика программы или конечная цель исследования выходит за рамки осведомленности эксперта, то процесс проведения РИ может оказаться слабо прогнозируемым, неэффективным или же даже безрезультатным. Таким образом, в данной предметной области существует научная проблема, имеющая вид следующего противоречия. С одной стороны, в информационной среде присутствует огромное количество разнородного программного обеспечения с МК, выполняемым на различных устройствах, имеющим различный формат хранения, целевое назначение, интерфейсы взаимодействия, условия функционирования и т. п. – то есть РИ необходимо применять для всего многообразия программного обеспечения. С другой стороны, отдельно взятые эксперты по РИ (или даже экспертные группы) знакомы лишь с рядом методов, имеют в наличии и используют определенный набор средств, восстанавливают из МК лишь определенную часть информации – то есть в рамках одного исследования РИ может применяться для ограниченного числа программного обеспечения. Первым же шагом для разрешения противоречия должно стать определение методологии (далее – Методология), как логической организации методов и средств проведения РИ. А поскольку (что будет показано далее) подобная Методология, обладающая необходимой полнотой, отсутствует, то будет предложена новая, основанная на огромном научно-практическом опыте автора по проведению РИ.

2. Обзор работ

Проведем обзор работ, посвященных вопросу РИ в части существующих Методологий его проведения, а также общих используемых методов и инструментария; при этом вопросы применения отдельных программных средств затрагиваться не будут, поскольку они являются частными решениями по автоматизации методов.

В [3] рассматриваются 3 информационно-технических процесса, связанные с РИ и решающие следующие задачи. Во-первых, РИ существующих физических объектов помогает создать их цифровые модели, переведя процесс исследования и разработки в информационную плоскость. Так, например, получение 3D-модели технических устройств с последующей ее модификацией позволяют не только проектировать инновационный опытный образец, но и производить его тиражированный выпуск. Во-вторых, РИ программного обеспечения необходим для определения принципов работы последнего, его архитектуры и алгоритмов, модификация которых позволяет расширять функциональные возможности. Частными задачами в рамках РИ указана расшифровка форматов файлов, восстановление исходного кода или потерянной метаинформации (имен классов, функций и переменных), построение графов потока управления и т. п. И, в-третьих, рассматривается обратная сторона РИ, связанная с защитой собственных программных разработок от РИ. Для этого предлагается применять алгоритмы шифрования и обфускации. Также рассматривается внедрение в собственные продукты недекларированных возможностей с целью контроля использования программ и для усложнения их РИ, что, впрочем, идет в разрез с законодательством РФ. Упоминаются и более специфичные способы защиты кода от РИ, такие как его мутация, рандомизация и использование защищенных сред выполнения. Однако, несмотря на некоторую проработанность области РИ, в статье отсутствует какая-либо полноценная Методология, имеющая не только формализованный вид, но и содержащая хотя бы какую-то последовательность методов (или шагов) и инструментов.

Работа [4] посвящена вопросу создания центра реверс-инжиниринга (далее – ЦРИ), актуальность которого (по мнению авторов) обосновывается «технологической блокадой» России со стороны «США и их несамодостаточных стран-сателлитов», что требует создания собственных технических решений на основе существующих путем сбора и переработки информации о них. Подчеркивается важность не дублирования готовых решений, а создания на их основе новых инновационных с применением всего аппарата научно-практической мысли (например, теории решения изобретательских задач). Предложен алгоритм работы такого

ЦРИ, основной последовательностью шагов которого является следующая: составление, утверждение и уточнение технического задания; структурный и функционально-стоимостной анализ проектного проекта, а также патентные исследования; испытания натурного образца (в случае необходимости); выбор и копирование необходимых элементов; модернизация конструкции и ее патентная защита; формирование и передача финальных документов. В качестве недостатков предложенного решения отмечены трудности в определении слабых мест исходного изделия и способах оптимизации процесса из РИ.

Целью исследования в [5] является оценка значимости патентной информации для обеспечения РИ, а задачей – выявление проблем, связанных с применением в РФ РИ к физическим объектам. В статье дискутируется терминология предметной области; так, под РИ (или реинжинирингом, обратной разработкой) понимается как процесс анализа продукта для сбора информации о нем, включая построение его 3D-модели, так и его воспроизведение. Приводятся два основных этапа РИ, а именно следующие: 1) сбор данных об объекте, его анализ и оценка правовых аспектов; 2) воспроизведение точной копии объекта, в том числе после его доработки. Подчеркивается обоснованность проведения РИ только при наличии материалов и средств для дальнейшего воспроизводства «клона». Делается вывод касательно важности поисково-исследовательской деятельности по сбору и анализу открытой патентной информации об объекте; приводится пример, что в патентах может быть найдено от 70 до 90 % технической информации об объекте.

В работе [6] описывается методика проведения РИ в программном продукте Siemens NX, позволяющая разрабатывать электронные макеты изделий перед его технологическим производством. В методике выделяются такие этапы, как импорт данных, совмещение и склейка фасетных тел, исправление ошибок геометрии, размещение и центрирование фасетного тела, получение сечений и/или цветовое выделение граней, создание модели и анализ ее точности, формирование «идеализированной» модели. Авторы указывают, что предложенная методика подойдет для всего спектра объектов любого масштаба и сложности.

Работа [7] обсуждает проблемы, существующие при практическом приложении РИ для производства ракетно-космической техники на основе существующих решений. К основным проблемам в части реализации на ЭВМ авторы относят следующие: дороговизна решения, высокая квалификация специалиста, слабая адаптивность моделей для всех разработчиков, рутинность процесса. Также отмечены второстепенные проблемы применения РИ для космических аппаратов, такие, как сложность управления проектом, нарушение геометрических связей, отсутствие стандартов РИ и несовместимость итоговых частей всей системы.

В статье [8] приводится мнение различных руководителей крупных IT-компаний касательно требований в вакансии к сотрудникам на позицию реверс-инженера, что, таким образом, может определять перечень знаний, умений, методов и средств, необходимых для проведения РИ. Все множество указанных требований может быть систематизировано в следующий перечень: мониторинг процессов в операционной системе (далее – ОС); практика проведения РИ; понимание основ работы аппаратного обеспечения и ОС; знание всего стека классических языков программирования (от ассемблерного до C/C++ и Python); работа с продуктом IDA Pro (включая плагин декомпиляции HexRays), WinDBG и OllyDbg, а также различными утилитами дизассемблирования и отладки.

Исследование [9] посвящено РИ сетевого протокола взаимодействия объектов. Для этого авторы делают обзор различных инструментов статического анализа трафика, что в итоге позволяет сформировать единый набор этапов проведения РИ, а именно следующих: первоначальная настройка параметров для выбора последующих методов, сбор и обработка данных, выделение в данных признаков для извлечения закономерностей внутри и между сетевыми сообщениями, идентификация сообщения для создания их семантических групп, определение формата сообщений для выделения их полей, раскрытие семантических значений полей, построение модели конечного автомата для состояний работы протокола, финальная пост-обработка результатов ручным или автоматическим способом.

Согласно обзору немногочисленного числа найденных релевантных работ, ни в одной из них не дается сколь-либо полноценного описания методологии РИ, а в самих исследованиях приводятся или достаточно общие этапы, или детализация их частных алгоритмов; при этом формализация РИ практически отсутствует. Также никак не рассматривается возможность применения для анализа МК машинного обучения, что с точки зрения автора является безусловным упущением [10, 11]. Таким образом, тема текущего авторского исследования является безусловно актуальной и новой.

3. Онтологическая модель

Перед описанием предлагаемой Методологии построим онтологическую модель проведения РИ, определяющую используемые далее основные понятия предметной области и их взаимосвязи; данная модель приведена на рисунке 1 (синим фоном отмечена основная деятельность РИ, серым – второстепенная деятельность, желтым – основной объект приложения, зеленым – основной результат, красным – побочные результаты).

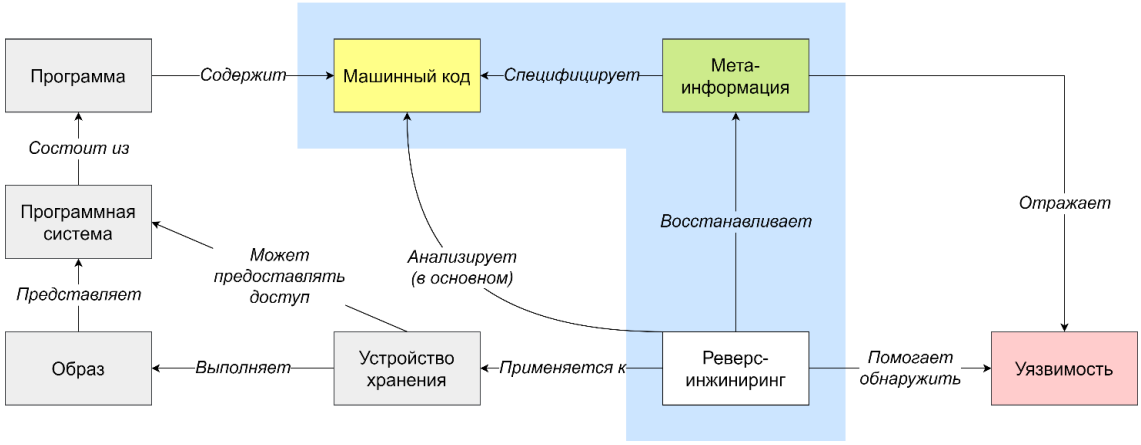


Рис. 1. Онтологическая модель проведения реверс-инжиниринга

Fig. 1. Ontological Model of the Reverse Engineering

Особенностью данной онтологической модели является то, что она сама по себе уже дает некоторое представление об основных принципах РИ.

На схеме присутствуют следующие элементы (их взаимосвязи отмечены курсивом):

1) Машинный код – низкоуровневое представление логики работы в виде инструкций, выполняемых процессором;

2) Программа – бинарный файл, *содержащий* экземпляр МК и информацию для его хранения и выполнения (например, имя файла, заголовок с информацией о процессоре выполнения и т. п.);

3) Программная система – логическая структура, *состоящая из* совокупности взаимосвязанных и взаимодействующих Программ (в вырожденном случае является единичной Программой), а также вспомогательных невыполняемых файлов;

4) Образ – *представление* Программной системы в виде монолитной бинарной сущности, как правило запакованной, сжатой или зашифрованной (например, в ISO-форме); с учетом специфики решаемых задач Образ может не всегда присутствовать в явном виде, как например в ОС Windows и Linux, которые собственными средствами *могут предоставлять доступ* к Программам;

5) Устройство хранения – техническое (реже, программное) устройство, *выполняющее* загружаемый в него Образ, который также может быть извлечен из устройства;

6) Метаинформация – сведения, *специфицирующие* МК и дающие полное человеко-ориентированное представление о его функционале (в основном включает псевдокод, алгоритмы, архитектуру, концептуальную модель, дополненные интерфейсами взаимодействия и т. п.) [12];

7) Реверс-инжиниринг – процесс восстановления Метаинформации о МК; в общем случае применяется к Устройству для извлечения и исследования Программ из Образа, хотя *основная* деятельность процесса состоит *в анализе* их МК [13];

8) Уязвимость – отличие итоговой реализации МК от задуманной или заявленной, что приводит к иному функционированию Программы, позволяющему реализовать угрозы (такое авторское введение понятия уязвимости раскрывается в [14, 15]); РИ *помогает обнаружить* Уязвимости, поскольку восстановленная в процессе этого Метаинформация *отражает* их признаки и более подходит для анализа экспертом.

4. Схема методологии

Предлагаемая Методология состоит из 4-х этапов, каждый из которых включает совокупность последовательно выполняемых шагов: *подготовительные мероприятия* (этап 1); *статическое исследование* (этап 2); *динамическое исследование* (этап 3); *документирование* (этап 4). При этом выходные результаты, полученные после применения одних шагов, являются входными данными для других шагов; исключением является 1-й шаг, имеющий на входе исходный МК, и шаги по описанию итоговых результатов, продуцирующие соответствующие финальные документы исследования. Шаги Методологии разбиты по этапам и имеют идентификаторы в следующем формате: *X.Y*, где *X* – номер этапа; *Y* – порядковый номер выполнения в рамках этапа.

Методология РИ в графическом виде имеет вид схемы, представленной на рисунке 2; на схеме используются следующие обозначения: прямоугольник с белым фоном – результат применения шага; круг с желтым фоном – идентификатор шага; сплошная стрелка – основное действие шага над результатом предыдущего; пунктирная стрелка – дополнительное использование шагов предыдущих результатов; пунктирный прямоугольник с белым фоном (и надписью «Этап 1. Подготовка объекта исследования») – шаги, описанные в текущей части цикла статей; пунктирный прямоугольник с серым фоном (и надписью «Этапы 2, 3. Анализ объекта исследования; Этап 4. Документирование») – шаги, которые будут описаны в следующих частях цикла.

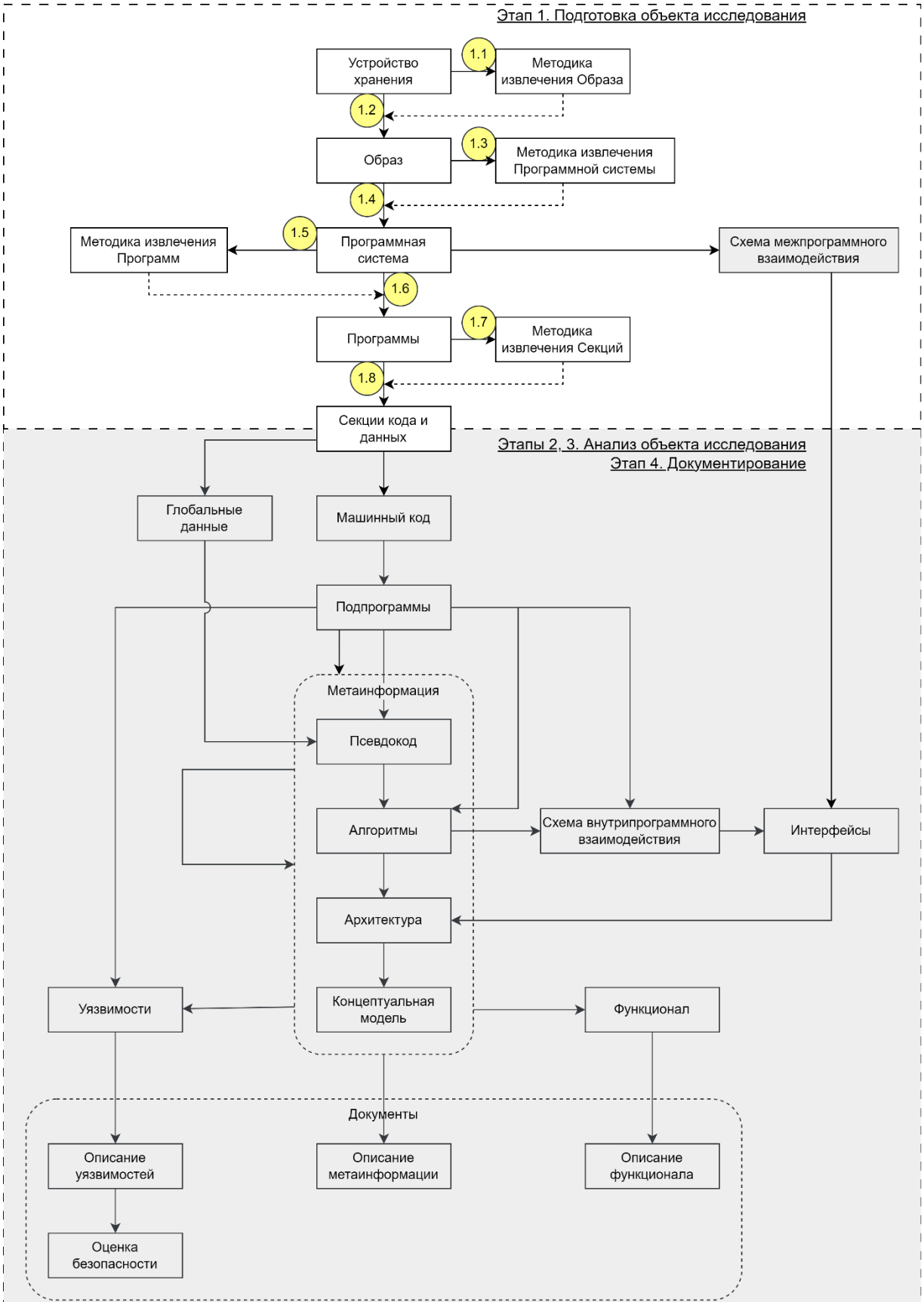


Рис. 2. Схема методологии реверс-инжиниринга (Этап 1)

Fig. 2. Reverse Engineering Methodology Diagram (Stage 1)

Согласно схеме (см. рисунок 2), в данной части статьи будут расписаны Шаги с 1.1 по 1.8; они полностью определяют 1-й этап РИ, предназначенный для извлечения из Устройства хранения основного объекта приложения РИ – программных секций с МК, а также дополняющих их секций с данными (которые носят вспомогательное назначение). Здесь и далее под секциями понимается деление файла программы на отдельные блоки данных определенного типа, таких, как следующие: «.text» – МК программы, «.data» – глобальные переменные, «rsrc» – ресурсы и др.

5. Этап 1. Подготовительные мероприятия

Шаги 1-го этапа Методологии, отвечающего за подготовку МК для непосредственного анализа, представлены далее.

Шаг 1.1. Сбор общей информации об устройстве хранения для разработки методики извлечения образа

Шаг является первоначальным для всей Методологии и текущего этапа, поскольку применяется к Устройству хранения, в котором содержится Образ с МК и всеми программно-логическими структурами. Способ получения такого Образа для дальнейшего анализа существенно зависит от конкретного типа устройства и механизмов хранения, что требует от экспертов изучение отдельных экземпляров устройств и создания соответствующих методик по работе с ними.

Формальная запись шага ($Step_{1.1}$) имеет следующий вид:

$$Method_{Image} = Step_{1.1}(Device),$$

где $Method_{Image}$ – методика получения Образа (*пер. на англ.* Image); $Device$ – исследуемое Устройство хранения.

Примером шага может быть сбор информации о материнской плате исследуемого компьютера (который в данном случае является Устройством хранения) с применением классической для этой задачи утилиты CPU-Z, пример графического окна которой показан на рисунке 3; так материнская плата имеет модель PRIME Z490-A производства ASUSTeK COMPUTER INC.

Затем, в соответствии с типом материнской платы, шаг может описать методику получения самого Образа. Так, в случае платы ASUS и ее UEFI-прошивки производства American Megatrends для этой задачи может быть применена утилита AfuWin, имеющая как графический, так и консольный вид.

Шаг 1.2. Применение к устройству хранения методики извлечения образа

Шаг заключается в применении методики (разработанной на Шаге 1.1), позволяющей получить Образ из конкретного экземпляра Устройства хранения.

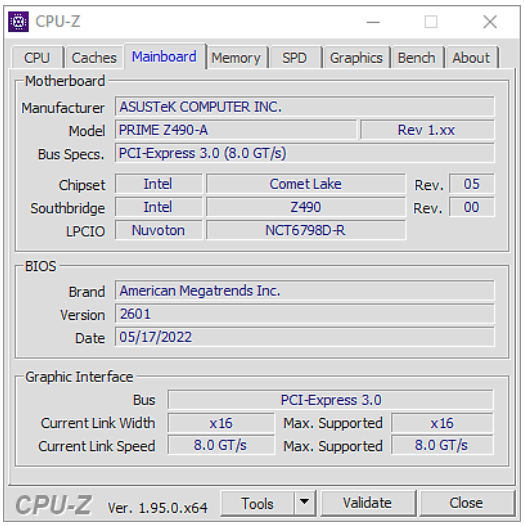


Рис. 3. Пример главного окна утилиты CPU-Z

Fig. 3. Example of the CPU-Z Utility Main Window

Формальная запись шага ($Step_{1.2}$) имеет следующий вид:

$$Image = Step_{1.2}(Method_{Image}, Device),$$

где $Image$ – Образ, полученный из исследуемого устройства ($Device$) с помощью специализированной для этого методики ($Method_{Image}$).

Примером шага может быть применение методики для получения Образа (в виде UEFI-прошивки) из компьютера с материнской платой ASUS, суть которой заключается в запуске утилиты AfuWin в графическом виде; главное окно утилиты представлено на рисунке 4.

Согласно рисунку 4, выбрано получение из материнской платы всех программных блоков, а нажатие кнопки Save приведет к сохранению Образа в бинарный файл на компьютере.

Шаг 1.3. Сбор общей информации об образе для разработки методики извлечения программной системы

Шаг применяется к полученному из Устройства хранения Образу и ставит своей задачей сбор информации о последнем для создания методики извлечения из него Программной системы, поскольку формат и внутренняя структура Образа может зависеть от производителя Устройства или же иметь различные механизмы защиты. Также, как было сказано ранее, в ряде случаев данный и последующий шаги можно пропустить, поскольку Программная система получается напрямую через Устройство; например, в случае Программ на персональном компьютере с типовой ОС.

Формальная запись шага ($Step_{1.3}$) имеет следующий вид:

$$Method_{ProgramSystem} = Step_{1.3}(Image),$$

где $Method_{ProgramSystem}$ – методика получения Программной системы (*пер. на англ.* Program Image); $Image$ – полученный ранее Образ.

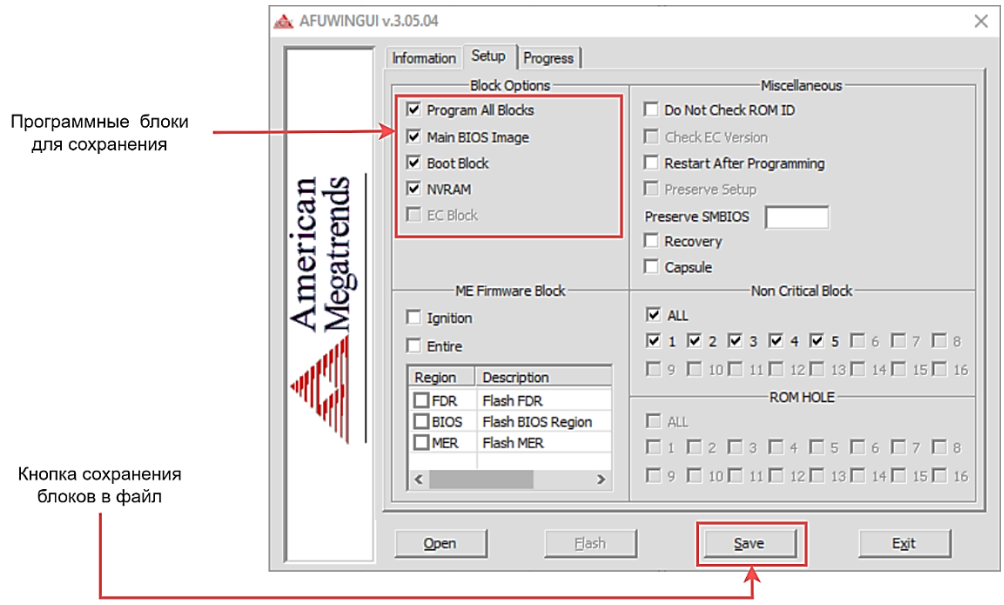


Рис. 4. Пример главного окна утилиты AfuWin

Fig. 4. Example of the AfuWin Utility Main Window

Примером шага может быть сбор информации об UEFI-прошивке (которая в данном случае является Образом) с применением достаточно редкого числа утилит, как стороннего, так и авторского производства. Так, для локализации файлов в Образе (которыми в случае UEFI-прошивки могут быть DXE-драйвера) возможно с применением утилиты UEFITool, пример главного окна которой (для версии NE alpha 67 от 20 июня 2023 г.) приведен на рисунке 5.

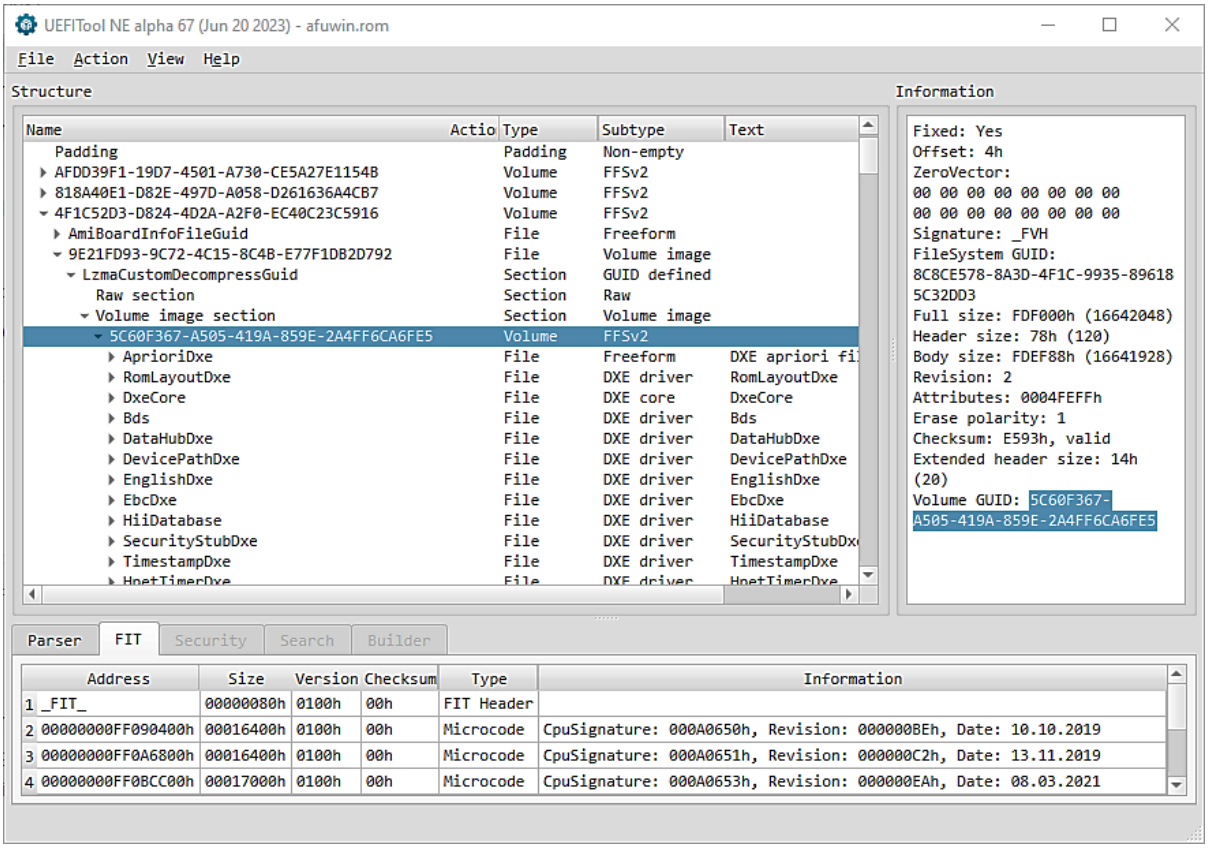


Рис. 5. Пример главного окна утилиты UEFI Tool (версии NE alpha 67)

Fig. 5. Example of the UEFI Tool Utility Main Window (version NE alpha 67)

Так, согласно рисунку 5, в Образе UEFI-прошивки присутствует том с GUID (*аббр. от англ.* Globally Unique Identifier, *пер. на русс.* Глобальный уникальный идентификатор) 5C60F367-A505-419A-859E-2A4FF6CA6FE5 и файловой системой FFSv2, содержащей набор драйверов – RomLayoutDxe, DxeCore, Bds и пр.

Затем, в соответствии с форматом Образа и способом его «раскрытия», результатом шага может стать описание методики извлечения Программной системы. В случае UEFI-прошивки для этой задачи подходит функционал UEFITool, позволяющий экспортировать бинарные данные всех логических элементов прошивки (как целых томов, так и отдельных драйверов) во внешние файлы.

Шаг 1.4. Применения к образу методики извлечения программной системы

Шаг заключается в применении методики (разработанной на Шаге 1.3), позволяющей получить Программную систему из конкретного экземпляра Образа.

Формальная запись шага ($Step_{1.4}$) имеет следующий вид:

$$ProgramSystem = Step_{1.4}(Method_{ProgramSystem}, Image),$$

где $ProgramSystem$ – Программная система, полученная из исследуемого Образа ($Image$) с помощью специализированной для этого методики ($Method_{ProgramSystem}$).

Примером шага может быть применение методики к UEFI-прошивке, суть которой заключается в ее открытии в утилите UEFITool, выборе нужного файлового тома и ручного применения к нему функционала утилиты по экспорту данных во внешние файлы. Как результат, можно получить бинарное представление Программной системы. В ряде случаев на данном шаге достаточно получить лишь список путей к файлам, поскольку их извлечение делается непосредственно из Образа, минуя работу с бинарным представлением Программной системы.

Шаг 1.5. Сбор общей информации о программной системе для разработки методики извлечения программ

Шаг применяется к полученной из Образа Программной системе и ставит своей задачей сбор информации о последней для создания методики извлечения из нее отдельных Программ, поскольку внутренняя структура системы и файлы с МК (из состава всех файлов) могут зависеть от ее формата и специфики исследуемого Устройства хранения.

Формальная запись шага ($Step_{1.5}$) имеет следующий вид:

$$Method_{Programs} = Step_{1.5}(ProgramSystem),$$

где $Method_{Programs}$ – методика получения Программ (*пер. на англ.* Programs); $ProgramSystem$ – полученная ранее Программная система.

Примером шага может быть сбор информации о файлах в томе UEF-прошивки для выделения DXE-драйверов (Программ, работающих на UEFI-фазе DXE, *аббр. от англ.* Driver Execution Environment, *пер. на русс*. Среда выполнения драйвера). В данном случае не все файлы в FFSv2 томах прошивок выполняемые, содержащие МК и, следовательно, выделение среди них Программ является отдельной задачей, которая может быть решена экспертно с применением упоминаемой утилиты UEFITool. Для этого можно воспользоваться стандартными названиями DXE-драйверов или изучением структуры файла, в котором будет присутствовать секция PE32, соответствующая PE-формату (*аббр. от англ.* Portable Executable, *пер. на русс.* Переносимый исполняемый) исполняемого файла для 32-битной разрядной системы. Одним из достаточно хорошо известных названий драйверов в UEFI является BDS (*аббр. от англ.* Boot Device Selection, *перев. на русс.* Выбор Загрузочного Устройства), реализующих одноименную фазу загрузки компьютера, а, следовательно, файл с таким именем скорее всего будет Программой. Пример отображения DXE-драйвера BDS с информацией о нем в утилите приведен на рисунке 6. Красным прямоугольником на белом фоне отмечена строка с названием DXE-драйвера, красным прямоугольником на синем фоне – его PE-секция с МК.

Все это позволяет сформировать итоговую методику получения Программ из Программной системы. Важно отметить, что в случае UEFI-прошивки на Шагах 1.3 и 1.4 достаточно лишь получить список файлов без непосредственного извлечения всей Программной системы в отдельный бинарный файл, поскольку получение конкретных Программ эффективнее делать непосредственно в утилитах по работе с UEFI-образом, такими, как UEFITool.

Шаг 1.6. Применение к программной системе методики извлечения программ

Шаг заключается в применении методики (разработанной на Шаге 1.5), позволяющей получить конкретные Программы из Программной системы. Формальная запись шага ($Step_{1.6}$) имеет следующий вид:

$$\begin{cases} Programs = \\ Step_{1.6}(Method_{Programs}, ProgramSystem), \\ Programs \equiv \{Program_i\} \end{cases}$$

где $Programs$ – множество Программ, полученных из исследуемой Программной системы ($Program\ System$) с помощью специализированной для этого методики ($Method_{Programs}$); $Program_i$ – i-я Программа из множества (для упрощения, индекс i далее опустим).

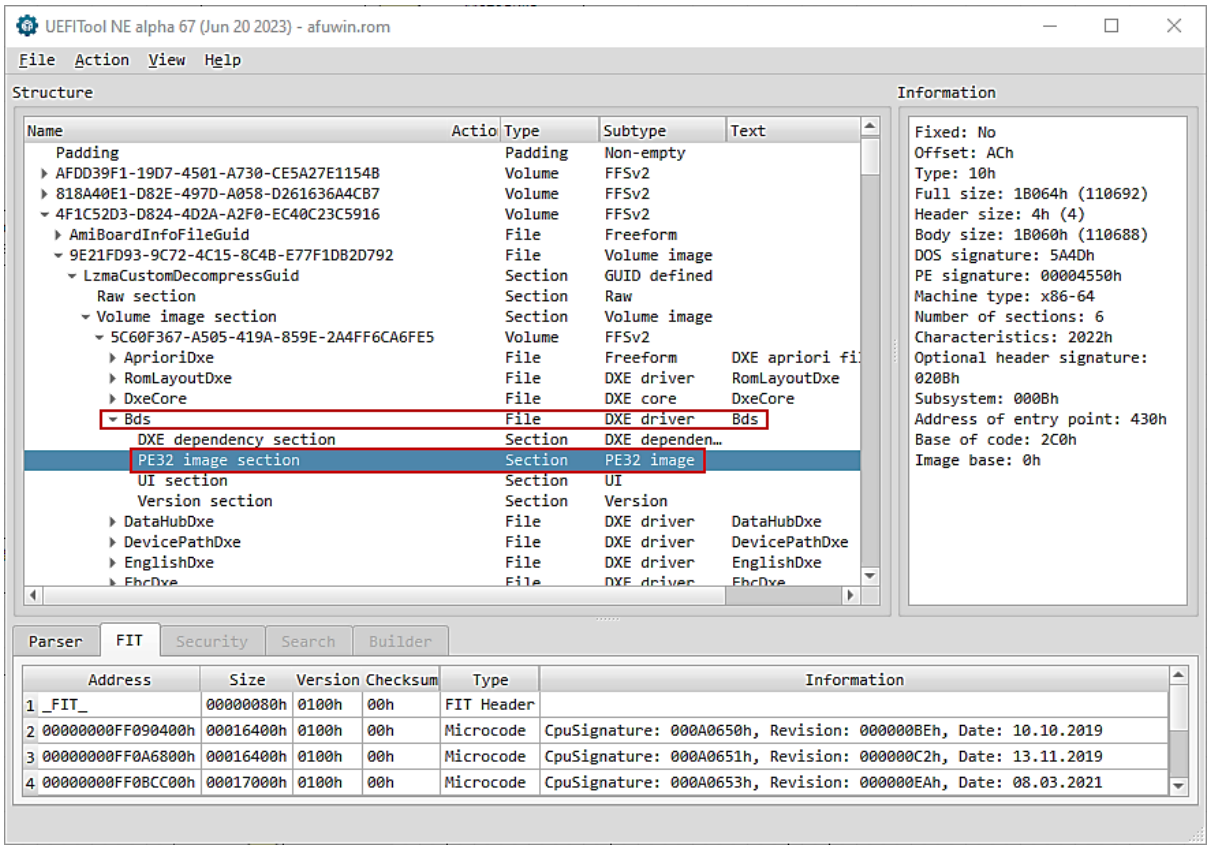


Рис. 6. Пример отображения информации о DXE-драйвере BDS в утилите UEFI Tool

Fig. 6. An Example of Displaying Information about the BDS DXE Driver in the UEFITool

Примером шага может быть применение методики к UEFI-прошивке, суть которой заключается в ее открытии в утилите UEFITool, выборе нужной Программы и применения к ней функционала утилиты по экспорту данных во внешний файл. Как результат, можно получить бинарное представление каждой Программы тома файловой системы; аналогичный функционал присутствует у авторской утилиты по работе UEFI-образом.

Шаг 1.7. Сбор общей информации о программах для разработки методики выделения секций

Шаг применяется к полученным из Программной системы Программам и ставит своей задачей сбор информации о каждой из последних для создания методики выделения секций с МК и используемыми им данными, поскольку внутренняя структура каждой Программы может зависеть от ее формата (примерами наиболее популярных являются PE для ОС семейства Windows и ELF для ОС семейства UNIX).

Формальная запись шага ($Step_{1.7}$) для каждой Программы имеет следующий вид:

$$Method_{Sections} = Step_{1.7}(Program),$$

где $Method_{Sections}$ – методика получения секций (*пер. на англ.* Sections); $Program$ – одна из полученных ранее Программ.

Примером шага может быть чтение первых байт Программы, в которых, как правило, расположена сигнатура, задающая формат файла: «MZ» (байты в шестнадцатеричном виде – 0x4D 0x5A) для PE-программы, «\x7fELF» (байты в шестнадцатеричном виде – 0x7f 0x45 0x4c 0x46) – для ELF-программ и др. Так, например, открытие DXE-драйвера для BDS в шестнадцатеричном онлайн-редакторе HexEd.it даст результат, представленный на рисунке 7. Первыми 2-мя байтами файла BdsDxe.efi являются 0x4D и 0x5A и, следовательно, файл представляет собой Программу в PE-формате.

Затем, в соответствии с форматом Программы, можно выбрать утилиту по работе с ней, которая позволяет извлекать программные секции – как правило она будет относиться к семейству дизассемблеров программ (т. е. с поддержкой их различных форматов). Одной из наиболее популярных такого рода утилит является продукт IDA Pro (*аббр. от англ.* Interactive DisAssembler, *пер. на русс.* Интерактивный ДизАссемблер), который позволяет не только определять форматы программ и процессора их выполнения, выделять в них секции, но и отображать МК, а также производить его извлечение автоматически с применением скриптов на С-подобном или Python языках [16]. Как альтернатива, существуют и специализированные программы по работе с заголовками файлов различных форматов.

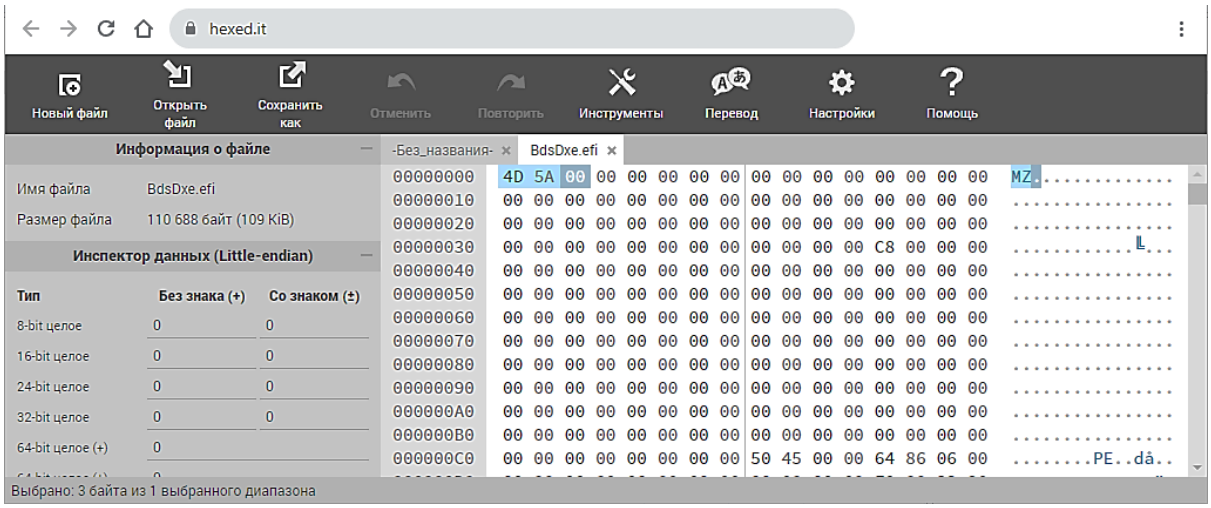


Рис. 7. Пример открытие DXE-драйвера для BDS в HexEd.it

Fig. 7. Example of Opening a DXE Driver for BDS in HexEd.it

Все это позволяет сформировать итоговую методику получения секций из каждой Программы в подходящей для этого утилите.

Шаг 1.8. Применение к программам методики выделения секций

Шаг является конечным для текущего этапа и заключается в применении методики (разработанной на Шаге 1.7), позволяющей получить секции с МК и данными в конкретной Программе. Формальная запись шага ($Step_{1.8}$) имеет следующий вид:

$$\begin{cases} Sections = Step_{1.8}(Method_{Sections}, Program) \\ Sections \equiv \langle Section_i \rangle \\ Class_{Section} \epsilon \{Section^{Code}, Section^{Data}\} \end{cases},$$

где $Sections$ – множество секций, полученных из исследуемой Программы ($Program$) с помощью специализированной для этого методики ($Method_{Sections}$); $Section_i$ – i-я секция из множества; $Class_{Section}$ – класс секции, которая содержит или код ($Section^{Code}$) или данные ($Section^{Data}$).

Примером шага может быть применение методики к DXE-драйверу с реализацией UEFI-фазы BDS – BdsDxe.efi, суть которой заключается в открытии файла в утилите IDA Pro и отображении окна Segments, пример которого приведен на рисунке 8. Так, Программа BdsDxe.efi была корректно открыта в утилите, а на экране отображены ее секции, одна из которых (с названием «.text») содержит МК.

Аналогичным образом применение утилиты PE Tools позволит вывести более детальную информацию о Программе, также включая и секции (пример вывода показан на рисунке 9).

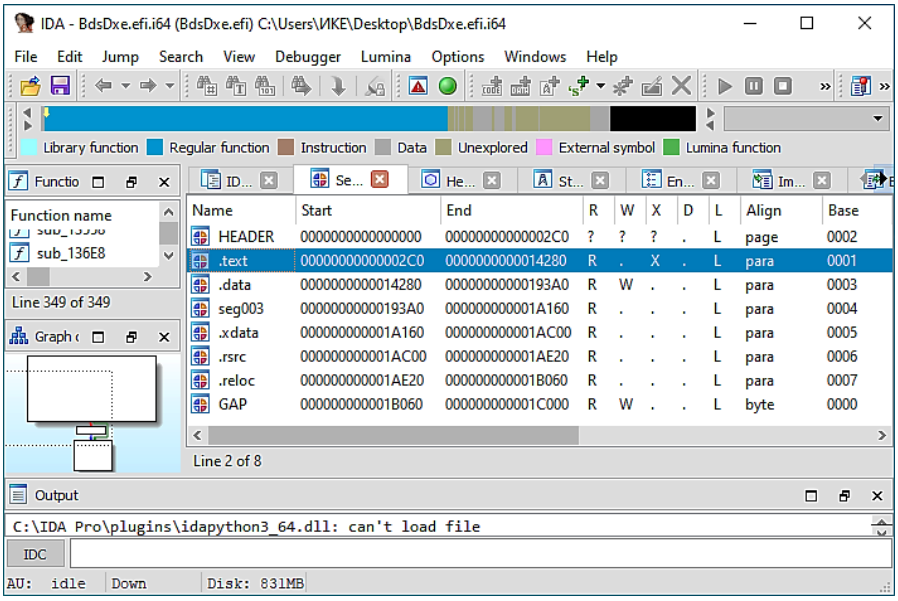


Рис. 8. Пример окна Segments утилиты IDA Pro

Fig. 8. Example of the IDA Pro Utility Segments Window

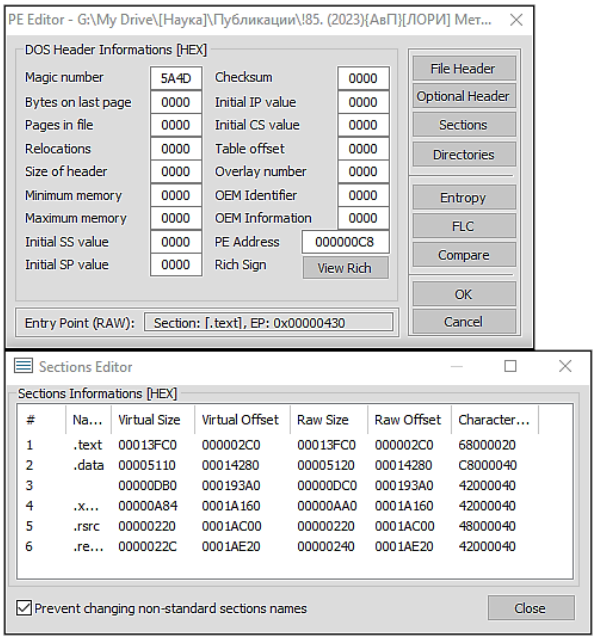


Рис. 9. Пример окон программы PE Tools

Fig. 9. Example of the PE Tools Utility Windows

Выделенные секции (называемые также сегментами) на рисунках 8 и 9 идентичны (за исключением специфичных для IDA Pro блоков данных HEADER и GAP, что подтверждает возможность применения обоих утилит.

Таким образом, после выполнения Шага 1.8 для анализа будут доступны программные секции, часть из которых содержит МК, а часть – используемые им данные, содержащие дополнительную информацию для работы шагов последующих этапов.

Заключение

Исследование посвящено построению единой Методологии проведения реверс-инжиниринга для машинного кода программ, основанной на существующих научных исследованиях предметной области и богатом авторском опыте. Обзор существующих работ показал практически полное отсутствие даже упоминания о каких-либо систематизированных и глубоко проработанных схемах в данной области. Как результат, предложена авторская методология, состоящая из 4-х этапов, шаги 1-го из которых описаны в первой (текущей) статье данного цикла, имеющая как схематичную форму, так и формализованную запись.

Новизной результата является системность подхода и масштаб охвата процесса по сравнению с существующим набором отдельно применяемых методов и средств, не всегда согласующихся друг с другом.

Теоретическая значимость результата заключается в создании полноценной схемы процесса (представляющей собой обобщенный алгоритм в виде операций и используемых данных), а практической – возможность реального применения от момента получения закрытого устройства до формирования итоговой документации с детальным описанием функционала машинного кода.

В следующих частях цикла статей будут расписаны шаги оставшихся 3-х этапов Методологии, сведенные также в единую таблицу с формой выполнения, а также входными и выходными информационными объектами.

Продолжение следует…

Список источников

1. Марков А.С., Цирлов В.Л. Опыт выявления уязвимостей в зарубежных программных продуктах // Вопросы кибербезопасности. 2013. № 1(1). С. 42–48.
2. Sabir U., Azam F., Haq S.U., Anwar M.W., Butt W.H., Amjad A. A Model Driven Reverse Engineering Framework for Generating High Level UML Models From Java Source Code // IEEE Access. 2019. Vol. 7. PP. 158931–158950. DOI:10.1109/ACCESS.2019.2950884
3. Баранова И.В., Батова М.М., Майоров С.В. Информационные инструменты реверсинжиниринга в стратегии деятельности инновационно-ориентированных структур // Теоретическая экономика. 2020. № 3(63). С. 28–35.
4. Передерий М.В. Реверс-инжиниринг в условиях инновационной инфраструктуры // Вестник Южно-Российского государственного технического университета (НПИ). Серия: Социально-экономические науки. 2015. № 5. С. 30–34.
5. Ивлиев Г.П., Эриванцева Т.Н. Патентная информация - источник ценных знаний для реинжиниринга // Право и цифровая экономика. 2022. № 3(17). С. 5–11. DOI:10.17803/2618-8198.2022.17.3.005-011
6. Нехорошев М.В. Методика реверс инжиниринга изделий в системе Siemens NX // Международная научно-техническая конференция «Проблемы и перспективы развития двигателестроения» (Самара, Россия, 23–25 июня 2021). Самара: Самарский национальный исследовательский университет имени академика С.П. Королева, 2021. Т. 1. С. 275–276.
7. Беляков А.А., Шулепов А.И. Проблемы практики реверс-инжиниринга космических аппаратов // Решетневские чтения: материалы XXV Международной научно-практической конференции, посвященной памяти генерального конструктора ракетно-космических систем академика М.Ф. Решетнева (Красноярск, Россия, 10–12 ноября 2021). Красноярск: Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, 2021. Ч. 1. С. 8–9.
8. Штомпель И. Вакансия: реверс-инженер // Системный администратор. 2014. № 11(144). С. 85–87.
9. Kleber S., Maile L., Kargl F. Survey of Protocol Reverse Engineering Algorithms: Decomposition of Tools for Static Traffic Analysis // IEEE Communications Surveys & Tutorials. 2019. Vol. 21. Iss. 1. PP. 526–561. DOI:10.1109/COMST.2018.2867544
10. Kotenko I., Izrailov K., Buinevich M. Static Analysis of Information Systems for IoT Cyber Security: A Survey of Machine Learning Approaches // Sensors. 2022. Vol. 22. Iss. 4. P. 1335. DOI:10.3390/s22041335

11. Израилов К.Е. Концепция генетической декомпиляции машинного кода телекоммуникационных устройств // Труды учебных заведений связи. 2021. Т. 7. № 4. С. 95–109. DOI:10.31854/1813-324X-2021-7-4-95-109
12. Kotenko, I., Izrailov, K., Buinevich, M., Saenko I., Shorey R. Modeling the Development of Energy Network Software, Taking into Account the Detection and Elimination of Vulnerabilities // Energies. 2023. Vol. 16. Iss. 13. PP. 5111. DOI:10.3390/en16135111
13. Долгова К.Н., Чернов А.В., Деревенец Е.О. Методы и алгоритмы восстановления программ на языке ассемблера в программы на языке высокого уровня // Проблемы информационной безопасности. Компьютерные системы. 2008. № 3. С. 54–68.
14. Израилов К.Е. Моделирование программы с уязвимостями с позиции эволюции ее представлений. Часть 1. Схема жизненного цикла // Труды учебных заведений связи. 2023. Т. 9. № 1. С. 75–93. DOI:10.31854/1813-324X-2023-9-1-75-93
15. Израилов К.Е. Моделирование программы с уязвимостями с позиции эволюции ее представлений. Часть 2. Аналитическая модель и эксперимент // Труды учебных заведений связи. 2023. Т. 9. № 2. С. 95–111. DOI:10.31854/1813-324X-2023-9-2-95-111
16. Ревнивых А.В., Велижанин А.С. Методика автоматизированного формирования структуры дизассемблированного листинга // Кибернетика и программирование. 2019. № 2. С. 1–16. DOI:10.25136/2306-4196.2019.2.28272

References

1. Markov A., Tsirlov V. Experience in Identifying Vulnerabilities in Software. *Voprosy kiberbezopasnosti*. 2013;1(1):42–48.
2. Sabir U., Azam F., Haq S.U., Anwar M.W., Butt W.H., Amjad A. A Model Driven Reverse Engineering Framework for Generating High Level UML Models From Java Source Code. *IEEE Access.* 2019;7158931-158950. DOI:10.1109/ACCESS.2019.2950884
3. Baranova I.V., Batova M.M., Mayorov S.V. Reverse Engineering Information Tools in the Strategy of Innovative-Oriented Structures. *Teoreticheskaia ekonomika*. 2020;3(63):28–35.
4. Perederiy M.V. Reverse Engineering in the Conditions of Innovation Infrastructure. *Bulletin of the South-Russian State Technical University (NPI). Series: Socio-Economic Sciences*. 2015;5:30–34.
5. Ivliev G.P., Erivantseva T.N. Patent Information as a Source of Valuable Knowledge for Reengineering. *Law and Digital Economy*. 2022;3(17):5–11. DOI:10.17803/2618-8198.2022.17.3.005-011
6. Nekhoroshev M.V. Reverse Engineering of Products in Siemens NX. *Proceedings of the International Scientific and Technical Conference on Problems and Prospects of Engine Building Development, 23–25 June 2021, Samara, Russia, vol.1*. Samara: Samara National Research University Publ.; 2021. p.275–276.
7. Belyakov A.A., Shulepov A.I. Problems of Spacecraft Reverse Engineering Practice Proceedings of the XXV International Scientific and Practical Conference Dedicated to the Memory of the General Designer of Rocket and Space Systems Academician M.F. Reshetnev, 10–12 November 2021, Krasnoyarsk, Russia. Part 1. Krasnoyarsk: Siberian State University of Science and Technology Publ.; 2021. p.8–9.
8. Shtompel I. Vacancy: Reverse Engineer. *Sistemnyi administrator*. 2014;11(144):85–87.
9. Kleber S., Maile L., Kargl F. Survey of Protocol Reverse Engineering Algorithms: Decomposition of Tools for Static Traffic Analysis. *IEEE Communications Surveys & Tutorials*. 2019;21(1):526–561. DOI:10.1109/COMST.2018.2867544
10. Kotenko I., Izrailov K., Buinevich M. Static Analysis of Information Systems for IoT Cyber Security: A Survey of Machine Learning Approaches. *Sensors*. 2022;22(4):1335. DOI:10.3390/s22041335
11. Izrailov K. The Genetic Decompilation Concept of the Telecommunication Devices Machine Code. *Proceedings of Telecommun. Univ*. 2021;7(4):95–109. DOI:10.31854/1813-324X-2021-7-4-95-109
12. Kotenko, I., Izrailov, K., Buinevich, M., Saenko I., Shorey R. Modeling the Development of Energy Network Software, Taking into Account the Detection and Elimination of Vulnerabilities. *Energies*. 2023;16(13):5111. DOI:10.3390/en16135111
13. Dolgova K.N., Chernov A.V., Derevenets E.O. Methods and Algorithms for Reconstructing Programs from Assembly to High Level Language. *Information Security Problems. Computer Systems*. 2008;3:54–68.
14. Izrailov K. Modeling a Program with Vulnerabilities in the Terms of Its Representations Evolution. Part 1. Life Cycle Scheme. *Proceedings of Telecommun. Univ*. 2023;9(1):75–93. DOI:10.31854/1813-324X-2023-9-1-75-93
15. Izrailov K. Modeling a Program with Vulnerabilities in the Terms of Its Representations Evolution. Part 2. Analytical Model and Experiment. *Proceedings of Telecommun. Univ*. 2023;9(2):95–111. DOI:10.31854/1813-324X-2023-9-2-95-111
16. Revnivykh A.V., Velizhanin A.S. Methods for Automated Formation of a Disassembled Listing Structure. *Cybernetics and programming*. 2019;2:1–16. DOI:10.25136/2306-4196.2019.2.28272

Статья поступила в редакцию 04.10.2023; одобрена после рецензирования 16.10.2023; принята к публикации 25.10.2023.
The article was submitted 04.10.2023; approved after reviewing 16.10.2023; accepted for publication 25.10.2023.

Информация об авторе:

ИЗРАИЛОВ Константин Евгеньевич

кандидат технических наук, доцент, старший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского Федерального исследовательского центра Российской академии наук

https://orcid.org/0000-0002-9412-5693

Обзорная статья
УДК 004.056
DOI:10.31854/1813-324X-2023-9-5-91-111

Обзор методов идентификации пользователя на основе цифровых отпечатков

Андрей Владимирович Осин, a.v.osin@mtuci.ru
Юрий Викторович Мурашко ✉, u.v.murashko@edu.mtuci.ru

Московский технический университет связи и информатики,
Москва, 111024, Российская Федерация

Аннотация: *Рассмотрены методы идентификации пользователей на основе цифровых отпечатков. Представлены основные подходы для формирования последних для браузера, который установлен на пользовательском устройстве и характеризует его принадлежность. Также описаны методы, применяемые для идентификации человека (пользователя) в процессе эксплуатации устройства. Представлены методы, использующие как динамику нажатий клавиш и взаимодействий с сенсорным экраном, голосовые и геолокационные данные, так и поведенческую биометрию и поведенческий профиль. В качестве развития подхода идентификации описана концепция непрерывной аутентификации. Приводится список общедоступных наборов данных, упоминаемых в рассмотренных в обзоре исследованиях, с указание ссылок для их скачивания. Приводится обширный список работ, отражающих современное состояние исследований в области цифровых отпечатков.*

Ключевые слова: *идентификация, цифровой отпечаток, браузер, поведенческая биометрия, непрерывная аутентификация*

Ссылка для цитирования: Осин А.В., Мурашко Ю.В. Обзор методов идентификации пользователя на основе цифровых отпечатков // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 91–111. DOI:10.31854/1813-324X-2023-9-5-91-111

A Review of User Identification Methods Based on Digital Fingerprint

Andrey Osin, a.v.osin@mtuci.ru
Yuri Murashko ✉, u.v.murashko@edu.mtuci.ru

Moscow Technical University of Communications and Informatics,
Moscow, 111024, Russian Federation

Abstract: *Methods of user identification based on digital fingerprints are considered. The main approaches for the browser fingerprints creation which is installed on the user's device and characterizes the device belonging to the user are presented. The methods used to identify a person (user) during the operation of the device are also described. Methods using both the dynamics of keystrokes and interactions with the touch screen, voice and geolocation data, as well as behavioral biometrics and behavioral profile are presented. The concept of continuous authentication is described as a development of the identification approach. A list of publicly available data sets mentioned in the studies reviewed in the review is provided, with links to download them.*

Keywords: *identification, digital fingerprint, browser, behavioral biometrics, continuous authentication*

For citation: Osin A., Murashko Y. A Review of User Identification Methods Based on Digital Fingerprint. *Proceedings of Telecommun. Univ.* 2023;9(5):91–111. DOI:10.31854/1813-324X-2023-9-5-91-111

Введение

Развитие средств вычислительной техники и связи позволило собирать, хранить, обрабатывать и передавать информацию в таких объемах и с такой оперативностью, которые были немыслимы раньше. Современный этап информатизации связан с повсеместным использованием персональной электронно-вычислительной техники, систем телекоммуникаций, сетей ЭВМ. Поэтому возрастает потребность в разработке и применении эффективных решений в сфере информационной безопасности. Под информационной безопасностью в общем виде следует понимать совокупность средств, методов и процессов (процедур), обеспечивающих защиту информационных активов и, следовательно, гарантирующих сохранение эффективности и практической полезности как технической инфраструктуры информационных систем, так и сведений, которые в таких системах хранятся и обрабатываются.

В данной статье рассматривается один из таких процессов обеспечения информационной безопасности – идентификация. Это процедура распознавания субъекта по его уникальному идентификатору, присвоенному ему ранее и занесенному в базу данных в момент его регистрации в качестве легального пользователя системы. В качестве распознаваемого субъекта может выступать физическое лицо, учетная запись компьютерной сети предприятия или, например, активность в трафике. Идентификатором субъекта могут выступать различные технологии, например, физические ключи доступа, биометрия, имя пользователя, файлы cookie, атрибуты трафика (IP-адрес, порт и др.).

Однако существуют методы фальсификации подобных идентификаторов, поэтому появляется необходимость в сборе данных для идентификации, где изменение одного или нескольких параметров не окажет большого воздействия на процесс идентификации. Для решения этой задачи может быть использована концепция цифрового отпечатка. Цифровой отпечаток – это информация, собранная, например, об удаленном устройстве для его идентификации. Поскольку данная концепция универсальная, то алгоритм на ее основе может быть применен во многих областях, например, для предотвращения мошенничества [1], формирования целевой рекламы [2], классификации вредоносного трафика [3], предотвращения обхода блокировок с использованием средств анонимизации [4], защита авторских прав [5–7] и др.

1. Идентификация на основе цифрового отпечатка браузера

Одна из самых распространенных областей, где используется цифровой отпечаток – идентификация браузера пользователя на основе собранных данных различными технологиями отслеживания при посещении какого-либо сайта. Необходимость использования именно цифрового отпечатка обусловлена тем, что такие идентификаторы, как файлы cookie, возможно обойти, а фальсифицировать весь массив данных, собранных о браузере, намного сложнее.

1.1. Исследования цифрового отпечатка браузера

В 2010 г. П. Экерсли из Electronic Frontier Foundation провел эксперимент с использованием программы Panopticlick [8]. Общаясь в социальных сетях и на популярных веб-сайтах, он за две недели собрал 470 161 цифровых отпечатков с данными из заголовков HTTP, JavaScript и плагинов, таких как Flash или Java. Из них 94,2 % были уникальными. Это сильно повлияло на конфиденциальность пользователей, так как браузер с редкими параметрами может быть легко идентифицирован в Интернете.

В 2012 г. в исследовании [9] была изучена возможность применения библиотеки для создания двухмерных изображений Canvas API при формировании цифрового отпечатка браузера. В исследовании отмечается, что обработка шрифтов может различаться в зависимости от устройства, поскольку операционная система, версия браузера, видеокарта, установленные шрифты, субпиксельные подсказки и сглаживание играют роль в создании окончательного видимого пользователем растрового изображения (рисунок 1). В ходе эксперимента они наблюдали 50 различных изображений из 300 образцов.

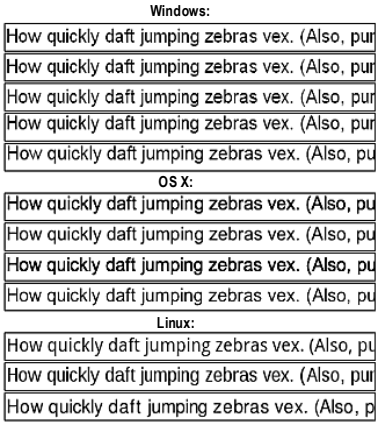


Рис. 1. Отрисовка текста с помощью Canvas API в разных операционных системах

Fig. 1. Drawing Text with Canvas API in Different Operating Systems

Также в [9] изучали использование WebGL API для получения дополнительных характеристик цифрового отпечатка браузера. Это библиотека для отображения интерактивных 3D-объектов в браузере и манипулирования ими с помощью JavaScript без необходимости использования плагинов.

В ходе эксперимента наблюдалось 50 различных отрисовок из 270 образцов, что объясняется разницей в аппаратном и программном обеспечении, при которой обработка не идентична на различных устройствах.

В 2017 г. в [10] разработан метод получения цифровых отпечатков, который также использует библиотеку WebGL при идентификации устройств. С помощью серии из 31 задачи отрисовки 3D-объектов тестируются параметры компьютерной графики для извлечения характеристик устройства, что позволило однозначно идентифицировать более 99 % из 1903 протестированных устройств.

В [11] описан еще один способ извлечения характеристик устройства для дополнения цифрового отпечатка браузер (рисунок 2). Способ основан на использовании Web Audio API, предоставляющий интерфейс для генерирования аудиосигнала и применения к нему специфических операций, таких как сжатие или фильтрация. Отмечается, что процесс получения цифровых отпечатков с использованием Web Audio API аналогичен процессу получения цифровых отпечатков с помощью Canvas API, поскольку обработанные сигналы будут иметь различия из-за программного и аппаратного обеспечения устройств.

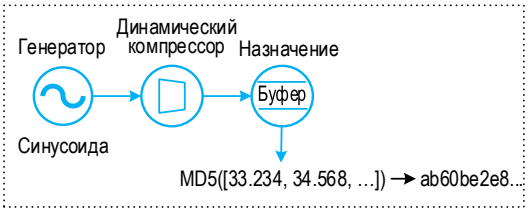


Рис. 2. Процесс извлечения характеристик с помощью Web Audio API

Fig. 2. Feature Extraction Process Using the Web Audio API

Современные браузеры предусматривают возможность разрабатывать собственные расширения, позволяющие увеличить его функциональные возможности. Однако из-за того, как расширения интегрируются в браузеры, некоторые из них можно обнаружить. Так, в исследовании [12] рассматривается механизм обнаружения расширений на основе доступности его ресурсов, например, логотипа. Поскольку доступ к таким ресурсам возможен в контексте любой веб-страницы, то можно использовать этот механизм для обнаружения наличия или отсутствия определенного расширения. Таким образом было обнаружено 12154 расширения браузера Chrome из 43429 и 1003 расширения браузера Firefox из 14896.

Похожее исследование было проведено в [13]. Авторы запрашивали ресурсы поддельных и существующих расширений и измеряли разницу во времени между вызовами (рисунок 3). Утверждается, что с помощью этого метода можно обнаружить любое расширение браузера. Авторы также провели исследование 204 пользователей с целью обнаружения 2000 расширений: они пришли к выводу, что 56,86 % пользователей уникальны.

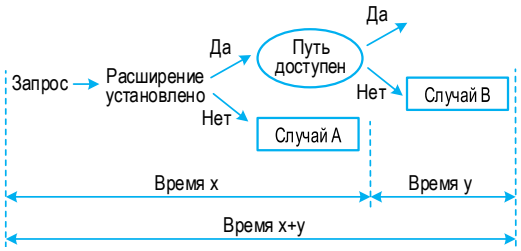


Рис. 3. Схема контроля доступности ресурсов

Fig. 3. Resource Availability Control Scheme

В [14] проведено исследование, которое заключается в выявлении побочных эффектов, вызываемых расширениями браузеров. Например, если расширение добавляет кнопку, то ее можно обнаружить путем анализа DOM (*аббр. от англ.* Document Object Model – объектная модель документа) вебстраницы (рисунок 4). В исследовании, основанном на 854 пользователях и обнаружении 1656 расширений, авторы пришли к выводу, что 14,10 % пользователей уникальны.

```
<div class="ProfileTweet-action action-pocket-container">
    <a class="js-tooltip"  href="#" data-original-title="
        Save to Pocket">
    <span class="icon icon-pocket"></span>
        </a></div>
<div class="ProfileTweet-action…
```

Рис. 4. Изменения DOM, внесенные расширением

Fig. 4. DOM Changes Made by the Extension

Еще один способ получить информацию об устройстве – сравнить возможности его центрального и графического процессоров. С помощью JavaScript можно запускать ряд задач и измерять время, необходимое для их выполнения. Однако самая большая трудность при использовании таких тестов производительности заключается в том, чтобы правильно оценить разницу и погрешность измерений. Два значения времени могут быть разными, потому что они были собраны с двух разных устройств, но они также могут принадлежать одному устройству, где появился новый фоновый процесс, нарушивший условия измерений.

В [15] использовалось 39 различных тестов для определения производительности «движка» Java Script; была показана возможность определения браузера и его версии с точностью 79,8 %. Но самым большим недостатком предложенного подхода является то, что для запуска полного набора тестов требуется в общей сложности 190,8 с. В отличие от большинства атрибутов, описанных выше, которые можно собрать за миллисекунды, эта разница во времени делает практически невозможным применение таких методов в реальных условиях.

В [16] используется библиотека WebGL API для отображения сложных 3D-сцен и измерения количества кадров, отображаемых браузером. Авторы показали, что оценка производительности графического процессора позволяет обнаружить различия между устройствами, поскольку небольшой графический процессор, например в смартфоне, будет значительно уступать в скорости выполнения операций новейшей графической карте высокого класса.

В [17] выполнен анализ конфиденциальности Battery Status API, где предоставляется такая информация об аккумуляторе устройства, как уровень его заряда, оставшееся время заряда и разряда, а также заряжается ли устройство в данный момент. Целью создания данной библиотеки было предоставление веб-разработчикам возможности реализации энергоэффективных приложений, однако в исследовании отмечается, что уровень заряда аккумулятора можно использовать в качестве краткосрочного идентификатора на сайтах, а повторные считывания могут помочь определить его емкость.

Важным аспектом цифрового отпечатка браузера является его эволюция с течением времени. Поскольку цифровой отпечаток является прямым отражением устройства пользователя и его окружения, он подвержен изменениям по мере модификации, настройки или обновления компонентов системы. Чтобы обеспечить долгосрочное отслеживание, необходимо понимать эти изменения и предвидеть, как может измениться цифровой отпечаток.

В [18] проведено обширное исследование, в ходе которого авторы ежедневно собирали цифровые отпечатки у добровольцев с помощью расширений браузеров Firefox и Chrome. По результатам исследования были выделены три типа эволюции цифровых отпечатков: автоматическая эволюция, вызванная обновлениями программного обеспечения; контекстно-зависимая эволюция, отражаемая изменениями в окружении пользователя; и эволюция, инициированная пользователем, вызванная изменением настроек браузера. Отмечается, что эволюция цифровых отпечатков браузера сильно зависит от типа устройства и того, как оно используется. Также авторы попытались со временем объединить цифровые отпечатки, принадлежащие одному и тому же устройству. Собирая цифровые отпечатки каждые три дня, их алгоритм мог отслеживать устройство в среднем 51,8 дня. Также удалось отследить 26 % устройств в течение более 100 дней, доказав, что снятие цифровых отпечатков браузера может эффективно использоваться в дополнение к другим методам идентификации.

Наиболее значимые работы в области цифровых отпечатков браузера приводятся в таблице 1.

ТАБЛИЦА 1. Основные исследования цифровых отпечатков браузера

TABLE 1. Basic Research on Digital Browser Fingerprints

Ссылка	Год	Цитируемость	Описание	Параметры
[8]	2010	1215	Electronic Frontier Foundation запустил веб-сайт, на котором посетители могут проверить цифровой отпечаток своего браузера. Собрав образцы из 470161 цифровых отпечатков, они измерили как минимум 18,1 бит энтропии, возможной при снятии отпечатков браузера.	• User Agent • HTTP ACCEPT headers • Cookie enabled? • Timezone • System fonts • Partial supercookie test
[9]	2012	406	Исследователи из Калифорнийского университета в Сан-Диего показали, как Canvas API и WebGL API можно использовать при создании цифровых отпечатков браузера.	• Canvas API • WebGL API
[11]	2016	666	Web Audio API является одним из последних дополнений в наборе инструментов получения цифровых отпечатков браузера. Обнаружены скрипты, которые обрабатывают звуковой сигнал для получения цифровых отпечатков браузера.	Web Audio API
[12]	2017	55	Исследования направлены на обнаружение расширений браузера по URL-адресам, изменению DOM веб-страницы и по разнице по времени между вызовами расширений.	Расширения браузера
[13]		89		
[14]		42		

1.2. Наборы данных

Обычно выделяют три крупномасштабных исследования по сбору цифровых отпечатков браузера, которые оказали большое влияние и способствовали развитию исследований в этой области: упомянутый ранее Panopticlick [8], AmIUnique [19] и Hiding in the Crowd [20]. В таблице 2 представлен краткий обзор этих трех масштабных исследований цифрового отпечатка браузера.

Эксперимент Panopticlick проводился в течение двух недель, в ходе которых было собрано около 470161 цифровых отпечатков браузера. Из них 83,6 % были уникальны, а если пользователи включали Flash или Java, то этот процент увеличивался

до 94,2 %. Список плагинов, список шрифтов и параметр user-agent были в то время наиболее информативными атрибутами. Также отмечается, что этот набор данных необъективен, поскольку данные получены от пользователей, которые заботятся о своей конфиденциальности в Интернете. Эти пользователи выполняли несколько простых мер, таких как ограничение файлов cookie или, возможно, использовали прокси-сервера для конфиденциального просмотра веб-страниц.

В эксперименте AmIUnique проведен анализ 118934 цифровых отпечатков браузера и выявлены новые результаты. Во-первых, подтверждены выводы исследования Panopticlick, поскольку 89,4 % собранных цифровых отпечатков браузера были уникальными. Однако за 6 лет, которые разделяют оба исследования, произошла эволюция различных атрибутов, составляющих цифровой отпечаток браузера. Если в начале десятилетия список плагинов и шрифтов были самыми информативными, то со временем сторонние плагины были отключены в основных браузерах из-за угрозы безопасности. В исследовании AmIUnique использовались полученные с помощью Canvas API параметры, энтропия которых была довольно высокой. В рамках эксперимента также были проанализированы различия между цифровыми отпечатками браузера на компьютерах и мобильных устройствах. Идентификация браузера на этих устройствах в значительной степени зависела от HTTP-заголовков и параметров, полученных с помощью Canvas API. В ходе анализа было установлено, что 81 % цифровых отпечатков браузера с мобильных устройств уникальны. Также, в исследовании обозначается, что такие простые изменения, как наличие стандартных HTTP-заголовков или удаление плагинов, снижают уникальность цифровых отпечатков браузера на компьютерах на 36 %.

В исследовании Hiding in the Crowd проанализировано 2067942 цифровых отпечатков браузера. Их результаты дают новый уровень понимания этой области, поскольку всего 33,6 % цифровых отпечатков браузера из их набора данных были уникальными. По сравнению с вышеупомянутыми исследованиями, это число в два-три раза меньше. Если рассматривать мобильные устройства, то разница еще больше: 18,5 % цифровых отпечатков браузера мобильных устройств были уникальными по сравнению с 81 % из исследования AmIUnique. Данное исследование подчеркивает важность процесса сбора данных. В прошлом цифровые отпечатки браузера собирались на веб-сайтах, ориентированных на посетителей, которые знают о конфиденциальности в Интернете или могут быть более осторожными, чем обычные пользователи. В данном случае данные собирались на коммерческом сайте, ориентированном на более глобальную аудиторию. Эта особенность набора данных в сочетании с очень большим количеством собранных цифровых отпечатков браузера являются ключом к пониманию различий в их уникальности. Также отмечается, что цифровые отпечатки браузера на компьютерах в основном уникальны из-за комбинации атрибутов, в то время как на мобильных устройствах присутствует атрибуты, имеющие уникальные значения. В таблице 3 приведена сводка атрибутов браузера вместе с их энтропией.

ТАБЛИЦА 2. Обзор исследований цифрового отпечатка браузера

TABLE 2. Overview of Browser Digital Fingerprint Studies

	Panopticlick (2010) [8]	AmIUnique (2016) [19]		Hiding in the Crowd (2018) [20]	
	Компьютер	Компьютер	Мобильный	Компьютер	Мобильный
Количество	470,161	105,829	13,105	1,816,776	251,166
Уникальность	94,2 %	89,4 %	81 %	35,7 %	18,5 %

ТАБЛИЦА 3. Атрибуты браузера и их энтропия в исследованиях цифровых отпечатков браузера

TABLE 3. Browser Attributes and Their Entropy in Studies of Digital Browser Fingerprint

Атрибуты	Panopticlick (2010)	AmIUnique (2016)	Hiding in the Crowd (2018)
	Энтропия		
Заголовок HTTP User-Agent	10,000	9,779	7,150
Заголовок HTTP Accept	–	1,383	0,729
Кодирование содержимого	–	1,534	0,382
Язык содержимого	–	5,918	2,716
Список расширений	15,400	11,060	9,485
Включены файлы cookie	0,353	0,253	0,000
Использование локального/сеансового хранилище	–	0,405	0,043

Атрибуты	Panopticlick (2010)	AmIUnique (2016)	Hiding in the Crowd (2018)
	Энтропия		
Часовой пояс	3,040	3,338	0,164
Разрешение экрана и глубина цвета	4,830	4,889	4,847
Список шрифтов	13,900	8,379	6,904
Список заголовков HTTP	–	4,198	1,783
Платформа	–	2,310	1,200
Включена функция «Не отслеживать»	–	0,944	1,919
Canvas API	–	8,278	8,546
WebGL Vendor	–	2,141	2,282
WebGL Renderer	–	3,406	5,541
Использование блокировщика рекламы	–	0,995	0,045

1.3. Применение цифровых отпечатков браузера

Цифровые отпечатки браузера могут применяться для различных целей, включая идентификацию пользователей и защиту от мошенничества. Однако следует отметить, что это может вызвать определенные проблемы, такие как нарушение конфиденциальности и приватности.

1.3.1. Отслеживание

Поскольку цифровые отпечатки браузера могут с высокой точностью идентифицировать устройства в Интернете, то последствия для конфиденциальности очень важны. Собирая цифровые отпечатки браузера на нескольких веб-сайтах, третья сторона может узнать пользователя и соотнести его активность в браузере внутри и между сессиями. При этом пользователь не может контролировать процесс отслеживания, поскольку он выполняются в фоновом режиме.

Если устройство имеет уникальный цифровой отпечаток, оно может быть идентифицировано в Интернете без использования других идентификаторов, таких как cookie файлы или IP-адрес. Пользователи, передающие свои сетевые пакеты через VPN (виртуальную частную сеть), особенно уязвимы для цифровых отпечатков браузера, поскольку VPN маскирует только IP-адрес.

В отсутствие cookie файлов цифровые отпечатки браузера можно использовать для отслеживания различных устройств, скрывающихся за одним и тем же IP-адресом.

1.3.2. Выявление уязвимостей устройства

Анализируя содержимое цифрового отпечатка браузера, злоумышленник может выявить потенциальные уязвимости, сопоставив список установленных на устройстве компонентов с базой данных распространенных уязвимостей. Затем он может разработать эффективный вредоносный код для конкретного устройства, заранее зная его уязвимости. Например, через свойство navigator.plugins можно узнать, работает ли на устройстве устаревшая версия плагина Flash, и если он не обновлен, то злоумышленник может удаленно выполнить вредоносный код на устройстве.

Компании Malwarebytes и GeoEdge в своей работе «Operation fingerprint» подробно описали, как рекламные кампании используют цифровые отпечатки браузеров для доставки вредоносных программ на уязвимые устройства [21]. Программный код цифровых отпечатков встраивается непосредственно в JavaScript поддельных рекламных объявлений и определяет, является ли устройство уязвимым или нет. Если да, то на устройстве будет показано объявление с вредоносным кодом, которое перенаправляет на набор инструментов для эксплуатации уязвимостей.

1.3.3. Повышение безопасности в Интернете

Выявление уязвимостей устройства с помощью цифровых отпечатков браузера можно применять и с целью их устранения. С помощью простого сканирования безопасности системные администраторы могут легко выявить устройства с устаревшими компонентами и быстро установить обновления.

Еще одно применение цифровых отпечатков браузера – повышение безопасности в Интернете путем проверки фактического содержания цифрового отпечатка. Поскольку между собранными атрибутами существует множество зависимостей, можно проверить, был ли отпечаток подделан или соответствует ли он устройству, которому предположительно принадлежит.

Одной из первых компаний, внедривших метод сбора цифровых отпечатков для предотвращения мошенничества в Интернете, была ThreatMetrix – компания по безопасности, специализирующаяся на проверке транзакций в Интернете [22]. Было отмечено, что мошенники меняют свой IP-адрес, удаляют cookie файлы, а программный код ботнетов произвольно изменяет атрибуты устройств. Другие компании по безопасности, такие как Imperva [23], MaxMind [24], HUMAN [25], IPQualityScore [26],

Radware [27] или Sift [28], также используют цифровые отпечатки браузера для обнаружения ботов и необычной активности.

Цифровой отпечаток браузера можно применять в дополнении к паролю для аутентификации в Интернете. Проверяя цифровой отпечаток браузера при входе в систему, можно блокировать несанкционированный доступ с новых и неизвестных устройств. Некоторые компании включают в свои продукты решения по снятию цифровых отпечатков браузера для усиления аутентификации. SecurAuth является поставщиком адаптивного решения для контроля доступа. Как часть их многофакторной аутентификации, они включают систему аутентификации на основе цифровых отпечатков устройств [29]. Другая компания под названием TransUnion имеет решение под названием TruValidate [30], которое объединяет сбор информации об устройстве в рамках многофакторной аутентификации.

2. Идентификация на основе цифрового отпечатка пользователя

Цифровой отпечаток браузера формируется из описанных ранее статических атрибутов, т. е. извлекаются однократно при подключении пользователя к веб-сайту [31]. Однако в условиях роста числа киберпреступлений идентификация на основе статических характеристик не обеспечивает достаточный уровень безопасности системе. Перехват сеанса и атака типа «человек посередине» – это всего лишь два примера потенциальных угроз, которыми может воспользоваться злоумышленник, чтобы выдать себя за легального пользователя системы, которая применяет статическую идентификацию [32]. Поэтому эксперты по безопасности в настоящее время рассматривают возможность внедрения динамической непрерывной идентификации. Непрерывная идентификация может быть выполнена с помощью поведенческой биометрии, которая является одним из наиболее перспективных решений этой проблемы. Поведенческие системы динамической идентификации в настоящее время внедряются в банках, государственных организациях и других учреждениях для обеспечения эффективной системы защиты от киберпреступлений [33].

Главное преимущество использования поведенческих характеристик заключается в возможности идентификации конкретного пользователя, а не его устройство, как, например, при использовании цифрового отпечатка браузера [34]. В зависимости от условий и поставленных задач поведенческие характеристики могут применятся по отдельности, однако объединение их в цифровой отпечаток пользователя позволит увеличить эффективность идентификации.

2.1. Исследования цифрового отпечатка пользователя

2.1.1. Динамика нажатия

Наиболее часто используемой поведенческой характеристикой является динамика нажатия клавиш клавиатуры, которую можно собрать, например, при вводе пользователем пароля. Многие существующие решения основаны на статистических данных о конкретных событиях, количестве их повторений во времени или определенных сочетаниях этих событий. Наиболее часто используемые характеристики включают: продолжительность нажатия клавиши и интервалы между нажатиями, скорость набора текста (среднее количество нажатий клавиш за заданное время), наложение определенных комбинаций клавиш, соотношение использования клавиш «Shift» или «Capslock» для ввода прописных/строчных букв, количество ошибок, методы исправления ошибок, и использование клавиш навигации (со стрелками) для перемещения курсора. В [35] показано, что всего две характеристики и простой классификатор могут обеспечить достаточно эффективную идентификацию пользователя.

Нажатие клавиши генерирует три основных события: событие опускания клавиши, событие отпускания клавиши и событие нажатия клавиши (происходит при вставке символа в текст) [36]. Перечисленные события используются для извлечения характеристик, разделенных на две группы: глобальные и временные.

Глобальные характеристики описывают общее поведение пользователя при наборе текста, такое как частота ошибок, удаление символов, использование клавиш «Shift», «Control», «Alt», общая скорость набора текста (нажатия клавиш или слов в минуту).

Временны́е характеристики относятся к стилю нажатия определенных клавиш или их комбинаций, на основе которых могут быть извлечены значения времени между (рисунок 5) [37]:

– отпусканием одной клавиши и нажатием следующей;

– нажатием и отпусканием одной и той же клавиши;

– нажатием одной клавиши и отпусканием следующей;

– нажатиями одной и следующей клавиш;

– отпусканием одной и следующей клавиш.

Идентификация пользователей – не единственное применение динамики нажатия клавиш. Эти свойства также могут быть применены в целях профилирования пользователей. За последнее десятилетие методы распознавания на основе нажатий клавиш привлекают все большее внимание в качестве инструмента компьютерной криминалистки в поведенческой биометрии. Исследования в этой

области чаще всего сосредоточены на определении возраста и пола человека, но также в [38] была создана модель машинного обучения для оценки уровня образования. Хотя точность такой модели составляла более 85 %, это решение имело ряд ограничений, в том числе высокую вычислительную сложность, которая связана с длительным временем обучения модели.

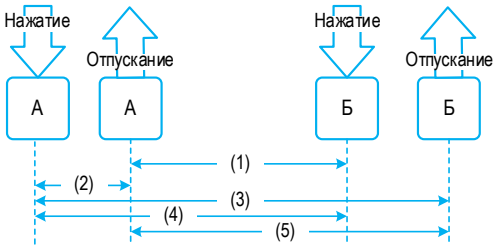


Рис. 5. Временны́е характеристики нажатия клавиш

Fig. 5. Time Characteristics of Key Presses

Тексты, используемые в экспериментах, могут иметь определенную длину (например, логин или пароль) или представлять собой спонтанное, ранее неопределенное содержание с неизвестным количеством букв. Первый вариант применяется в большинстве исследований, касающихся динамики нажатия клавиш. В настоящее время работы, проводимые в этой области, сосредоточены на усовершенствовании известных методов обработки временных характеристик [39–40]. Второй подход с нефиксированной длинной текста может быть использован для непрерывной идентификации. Этот подход исследуется как для стандартной клавиатуры, так и для смартфонов. В [41] авторы разработали Android-приложение IProfile для сбора событий нажатия клавиш на виртуальной клавиатуре. Во время эксперимента, в ходе которого участники набирали пароль, было извлечено 155 параметров. При этом точность различными методами классификации составила примерно 97 %.

2.1.2. Динамика движения курсора мыши и пальца

Другим примером поведенческой биометрии является динамика движения мыши, где идентификация пользователей осуществляется на основе того, как они используют свою мышь на компьютере. Поведенческий профиль создается путем извлечения характеристик, связанных с движениями мыши пользователя.

Ранние исследования динамики движения мыши были сосредоточены на распознавании электронных подписей пользователей. Например, в [42] использовали нейронные сети для изучения подписей, сделанных от руки. В 2003 г. в работе [43] проведено исследование подписывания с помощью мыши.

В 2007 г. в [44] представлено исследование возможности использования данных о перемещении мыши для идентификации личности. В предложенном подходе, актуальном и по сей день, выделяют следующие категории действия мыши:

– движение мыши;

– действие начинается с нажатия кнопки мыши, перемещения, а затем отпускания кнопки мыши;

– движение мыши, за которым следует щелчок или двойной щелчок;

– отсутствие движения.

Каждое такое действие состоит из следующих значений: пройденное расстояние в пикселях, прошедшее время в секундах и направление движения. Рассматриваются восемь направлений, охватывающих набор движений мыши в пределах 45 ° (рисунок 6).

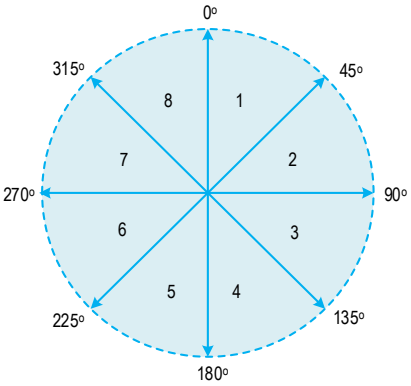


Рис. 6. Направления движения мыши

Fig. 6. Directions of Mouse Movement

Характеристики динамики мыши могут быть описаны набором параметров, полученных в результате анализа зафиксированных действий мыши. Эти параметры представляют собой составные части сигнатуры динамики мыши для конкретного пользователя, которая может быть использована для его идентификации. Среди них – средняя скорость движения каждого пройденного расстояния, движения в каждом направлении, среднее пройденное расстояние за определенный период времени по различным направлениям движения.

Аналогично динамике нажатия, современные мобильные устройства также привносят в динамику движения пальца по сенсору такие дополнительные параметры, как давление на экран и площадь экрана, на которую нажимает палец. Так, например, в [45] исследована точность жестов смахивания, используя продолжительность касания, длину траектории смахивания, среднюю скорость, ускорение, давление и площадь пальцев при движении. В рамках исследования было собрано 58 образцов от 40 испытуемых, и была проведена классификация, которая обеспечила уровень ошибок 0,004 %.

Также отмечается, что сенсорные данные зависят от приложений, поэтому не следует ожидать универсальной оценки эффективности для всех приложений, а вместо этого следует рассмотреть

варианты реализации в контексте конкретного приложения. Так, в [46] предлагается подход, суть которого заключается в исследовании преимуществ и недостатков идентификации на уровне устройств и приложений. Для эксперимента авторы разработали четыре приложения для 32 пользователей, которые использовали их в естественных условиях в течение десяти недель. Записывались характеристики сенсорного экрана, включая такие показатели, как координаты точки касания, давление пальца, площадь, ориентация экрана. В результате было отмечено, что показатели различаются в зависимости от приложения, следовательно, значимость характеристик различается в зависимости от приложения. Результаты экспериментов также показывают, что этот подход является более точным, чем подход, ориентированный на устройство, поскольку количество недостоверных признаков уменьшается.

2.1.3. Движение устройства

Мобильные устройства позволяют определять его движение с помощью встроенных акселерометров и гироскопов. Акселерометр измеряет ускорение в трех ортогональных пространственных измерениях: *x*, *y* и *z* [47]. Гироскоп измеряет вращение вокруг каждой из этих осей [48]. Комбинация этих измерений предоставляет пространство признаков, позволяющее моделировать движение пользователя.

Например, в [49] предлагается реализация, которая включает в себя так называемую воздушную подпись. Пока пользователь держит мобильное устройство в руках, он двигает его в воздухе, а измерения акселерометра записываются. Такая реализация требует от пользователя запуска приложения для сбора данных, поэтому она не является ни скрытой, ни прозрачной. Кроме того, сопоставление выполняется на сервере, что представляет угрозу безопасности в случае перехвата трафика. Тем не менее, алгоритм был протестирован на десяти добровольцах и показал 1,46 % ошибок 2 рода и 6,87 % ошибок 1 рода. В работе [50] предлагаются подобные попытки оценивания жестов махания, в [51] – жестов в свободной форме и в [52] – движения «поднятия руки» (т. е. доставания устройства из кармана, поднятия руки и поднесения телефона к уху).

Распознавание походки – это идентификация человека по тому, как он ходит, на основе методов машинного зрения, датчиков положения на полу или носимых датчиков. К последним можно отнести акселерометр и гироскоп мобильного устройства. Так, в [53] предлагается одна из таких систем идентификации пользователя, в которой продемонстрированно, как движения рук человека при ходьбе могут быть распознаны с помощью акселерометра и гироскопа в носимых устройствах. Для идентификации пользователя система применяет метод оценки и выбора признаков на основе корреляции, а также классификатор на основе скользящего окна. В результате авторы заявляют, что такая система отвечает ряду ключевых требований к идентификации пользователей по походке на устройствах с ограниченными ресурсами, таких как классификация в реальном времени, высокая точность идентификации и небольшое количество датчиков. Однако эта система ориентирована только на идентификацию по походке.

2.1.4. Голос

Метод на основе распознавания голоса объединяет физиологические и поведенческие характеристики для идентификации пользователя по его речи. Анатомические аспекты, например, голосовые связки, в сочетании с поведенческими характеристиками, такими как возраст или тональность, позволяют оценить множество характеристик, которые могут быть проанализированы статистически.

В [54] выделяется два способа распознавания голоса: текстозависимый и текстонезависимый. В первом случае пользователей просят произнести заранее определенную фразу, поэтому они знают о биометрической системе (следовательно, это не скрытая система). Из-за использования фиксированной фразы система работает точнее. Во втором – система пытается распознать говорящего независимо от того, что он произносит. Текстонезависимые системы полезны, когда существует меньший контроль над вводом, например, когда пользователь не знает о биометрической системе, что впоследствии обеспечивает большую гибкость. Однако, достижение высокой точности при распознавании голоса в такой системе является более сложной задачей из-за непредсказуемости говорящего.

Распознавание пользователя по особенностям голоса происходит по типичной схеме, начиная со сбора и предварительной обработки данных, извлечения и отбора характеристик и заканчивая моделированием и распознаванием образов. Как и в обычных системах на основе машинного обучения, качество характеристик вносит значительный вклад в точность идентификации пользователя. К таковым относятся краткосрочные спектральные, временные и ритмические, лингвистические характеристики, характеристики источника голоса и уровня разговора [55]. Краткосрочные спектральные характеристики представляют собой резонансные характеристики голосового тракта и часто извлекаются с высокой частотой для временны́х интервалов от 20 до 30 мс. Лингвистические и временные характеристики включают в себя интонационные и ритмические шаблоны, извлекаемые из длительных временных интервалов. Признаки разговорного уровня – это высокоуровневые свойства,

извлекаемые из текстового содержания устной речи, такие как частота слов или фраз. Качество характеристик измеряется их различительной способностью и устойчивостью к возможным внешним шумам (например, состояние пользователя и окружающая среда). Исследование [54] показало, что спектральные признаки обеспечивают высококачественное, простое и дискриминационное пространство признаков.

Последние достижения в области глубокого обучения привели к появлению инструментов синтеза требуемого голоса, которые создают синтетическую речь, произносимую голосом идентифицируемого пользователя, инструментов преобразования текста в речь, преобразующих произвольный текст в произносимые слова [56–57], а также инструментов, которые преобразуют существующие образцы голоса в тот же текст, который произносит идентифицируемый пользователь [58–60]. В [61] выделяют следующие методы синтеза голоса, позволяющие обойти системы идентификации на основе голоса: запись и воспроизведение [62–63], имитация голоса [64–65], машинный синтез (классический) [66–67], машинный синтез (на основе глубоких нейронных сетей) [68]. Для защиты систем распознавания речи от атак с использованием синтетической речи было предложено множество средств защиты. В то время как большинство из них сосредоточено на обнаружении синтетической речи [69–70], в [71] предлагается новое направление защиты – предотвращение несанкционированного синтеза голоса путем встраивания возмущений в аудиообразцы для их смещения в пространстве признаков.

2.1.5. Поведенческий профиль

Под поведенческим профилем понимаются интерактивные данные об использовании устройства, например, о том, как пользователь взаимодействует с мобильным устройством при совершении телефонных звонков, отправке текстовых сообщений и использовании приложений.

Работы в отношении интерактивных данных в основном связаны с обработкой нечисловых данных простым способом, например, с помощью подсчета частоты и категориальных представлений. В основном такой подход связан с невозможностью математической обработки этих значений. Значения интерактивных данных обычно представляют названия открытых или закрытых приложений или сетей Wi-Fi, к которым пользователь подключает устройство для доступа в Интернет [72].

Характеристики классифицируются следующим образом.

Категориальные: категориальные представления характеристик группируют интерактивные значения данных. Например, вместо того, чтобы перечислять названия каждого приложения для социальных сетей, вектор характеристик может просто включать само слово «социальная сеть» в качестве характеристики. Таким образом, если пользователь посещает пять социальных сетей, вектор характеристик будет указывать только на то, что приложения были для социальных сетей, не предоставляя точных сведений о них. В [73] рассмотрен этот подход для классификации трафика приложений. Время сбора данных группируется по времени суток, значения перемещения группируются по скорости, а значения местоположения устройства по наиболее частому и долгому пребыванию пользователя, например, на работе. Аналогично, в [74] классифицируется трафик приложений по таким группам, как музыка, обмен сообщениями и настройки. Очевидно, что такой подход сокращает значения характеристик до меньшего набора, который может быть слишком обобщенным, чтобы идентифицировать пользователей.

Частотные: представление характеристик на основе частоты основано на подсчете того, сколько раз выполняется то или иное действие [75]. Так, например, в [76] рассматривается частотное повторение действий для приложений и просмотра веб-страниц.

Последовательности: последовательности действий также являются характеристиками. В [77] предполагается, что порядок выполнения действий является значимым и уникальным для каждого пользователя. Этот подход похож на представление *n*-грамм для текста в приложениях стилометрии, где классификация основывается на частых *n*-кортежах, встречающихся в документах, идентифицируя стиль автора.

Для профилирования поведения были рассмотрены различные варианты реализации. Например, в [78] исследуется профилирование поведения как средство ассоциации условий с поведением. Пятьдесят студентов были привлечены к этому исследованию на один месяц, в течение которого записывались данные GPS, GSM, системные данные, журнал вызовов, данные датчиков и взаимодействия. Все добровольцы смогли подтвердить, что 95 % ассоциаций, обнаруженных исследователями, были верными.

Поведенческое профилирование также может использоваться для классификации вредоносных программ, так, в [79] предложено приложение Andro-prolifer как средство анализа системных вызовов, их аргументов и системных журналов для обнаружения вредоносных приложений на мобильных устройствах. В этой работе предпринята инновационная попытка устранить недостатки предыдущих подходов к обнаружению вредоносного программного обеспечения, которые в основном фокусировались на частоте системных вызовов, поскольку количество вызовов обычно невелико.

Авторы используют базу данных из 709 образцов вредоносного программного обеспечения и 350 доброкачественных образцов, достигая средней точности классификации 99 %. Наиболее значимые работы в области цифровых отпечатков пользователя приводятся в таблице 4.

ТАБЛИЦА 4. Основные исследования поведенческой биометрии

TABLE 4. Basic Research on Behavioral Biometrics

Ссылка	Год	Цитируемость	Описание	Параметры
[37]	2013	391	Представлен подробный обзор и сравнение исследований в области биометрии динамики нажатия клавиш.	• время между отпусканием одной клавиши и нажатием следующей; • время между нажатием и отпусканием одной и той же клавиши; • время между нажатием одной клавиши и отпусканием следующей; • время между нажатием одной и следующей клавиш; • время между отпусканием одной и следующей клавиш.
[44]	2007	480	Представлена новая форма поведенческой биометрии, основанная на динамике мыши, которая может быть использована в различных приложениях безопасности.	• средняя скорость движения; • средняя скорость движения в каждом направлении; • среднее пройденное расстояние за определенный период времени по различным направлениям движения.
[47]	2016	34	Исследования направлены на изучение возможности идентификации пользователей по значениям с датчиков акселерометра и гироскопа.	Координаты ускорения в трех ортогональных пространственных измерениях и вращения вокруг каждой из них.
[48]		17		
[54]	2002	933	Представлен краткий обзор методов распознавания речи, описывая применения, основные методы и некоторые показатели эффективности. Приведены методы синтетической речи и предотвращения несанкционированного синтеза голоса.	• краткосрочные спектральные характеристики; • временные и ритмические характеристики; • характеристики источника голоса; • лингвистические характеристики; • характеристики уровня разговора.
[55]	2016	632		
[69]	2020	49		
[70]	2019	58		
[71]	2021	28		
[57]	2013	22	В работах выделены разновидности интерактивных данных о поведении пользователей мобильных устройств.	• категориальные характеристики: тип приложения, время суток, скорость перемещения, местоположение; • частотные характеристики: частота запуска приложения, частота действий в приложении, частота просмотра веб-страниц; • характеристики последовательностей: порядок запуска приложений, порядок действия в приложении, последовательность символов в набираемом тексте.
[60]	2016	232		
[61]	2013	156		

2.2. Наборы данных

Существует множество наборов данных, однако многих из них нет в свободном доступе или содержат малое количество атрибутов. Поэтому для исследовательских целей стоит выделить 3 общедоступных крупномасштабных набора данных: UMDAA-02 [80], HuMIdb [81] и BehavePassDB [82]. В таблице 5 представлен краткий обзор этих трех исследований по поведенческой биометрии. Набор данных UMDAA-02 состоит из 141,14 ГБ сигналов датчиков смартфона, собранных у 48 добровольцев на телефонах Nexus 5 в течение двух месяцев [83]. К датчикам сбора данных относится фронтальная камера, сенсорный экран, гироскоп, акселерометр, магнитометр, датчик освещенности, GPS, Bluetooth, WiFi, датчики приближения, температуры и давления. Приложение для сбора данных также сохраняло время событий блокировки и разблокировки экрана, временные метки начала и окончания звонков, текущее открытое на экране приложение и т. д.

ТАБЛИЦА 5. Обзор исследований по поведенческой биометрии

TABLE 5. Overview of Behavioral Biometrics Studies

Объекты исследования	UMDAA-02 (2016), шт	HuMIdb (2020), шт	BehavePassDB (2022), шт
Сенсоры	10	14	15
Пользователи	48	600	81
Устройства	–	600	81

Набор данных Human Mobile Interaction (HuMIdb) включает более 5 ГБ данных для широкого спектра мобильных датчиков, полученных без сценария [84]. В базе данных хранятся данные от 14 датчиков

во время естественного взаимодействия человека с мобильным устройством, осуществляемого 600 пользователями. Для сбора данных было создано Android-приложение, которое собирает данные от датчиков во время выполнения пользователями 8 простых задач с помощью собственных мобильных устройств без какого-либо контроля (т. е. пользователи могли стоять, сидеть, ходить, находиться в помещении, на улице, днем или ночью и т. д.). Участники для эксперимента были отобраны по всему миру, что позволило получить более разнообразных участников, чем в предыдущих исследованиях.

BehavePassDB включает данные, также полученные во время естественного взаимодействия человека с мобильным устройством [85]. Сбор данных проводился в течение четырех сессий, каждая из которых была разделена по крайней мере 24-часовым промежутком. Испытуемых просили установить Android-приложение на свой смартфон и выполнить восемь заданий в неконтролируемом сценарии. Процесс сбора данных был разработан таким образом, чтобы имитировать наиболее важные сценарии взаимодействия человека и устройства в рамках прозрачной непрерывной аутентификации, т. е. поведенческие биометрические характеристики пользователя постоянно проверяются в течение всего времени использования устройства, не прерывая процесса идентификации. Кроме того, сессии сбора данных сбыли структурированы таким образом, чтобы сбалансировать количество собираемой информации и легкость сбора, чтобы вовлечь большое количество пользователей.

В таблице 6 приведена сводка атрибутов, собираемых в исследованиях по поведенческой биометрии.

ТАБЛИЦА 6. Информации и ее источник в исследованиях по поведенческой биометрии

TABLE 6. Information and Its Source in Behavioral Biometrics Research

Источник информации	UMDAA-02	HuMIdb	BehavePassDB
	Информация		
Акселерометр	*x, y, z*	*x, y, z*	*x, y, z*
Гироскоп	*x, y, z*	*x, y, z*	*x, y, z*
Сенсор гравитации	*x, y, z*	Значение ускорения	*x, y, z*
Датчик освещенности	Значение освещенности	Значение освещенности	Значение освещенности
Линейный акселерометр	*x, y, z*	*x, y, z*	*x, y, z*
Магнитометр	*x, y, z*	*x, y, z*	*x, y, z*
Касание экрана	*x, y*, давление, тип действия	*x, y*, давление, тип действия	*x, y*, тип действия
Разрешение экрана	*x, y*	–	–
Нажатие клавиши	ASCII-код, давление, тип	ASCII-код, давление	ASCII-код
Ориентация устройства	–	Значение ориентации (альбомная или книжная)	Значение ориентации (альбомная или книжная)
Датчик давления	–	–	Значение давления
Датчик приближения	–	Уровень приближения	Уровень приближения
GPS	Широта, долгота, точность	Широта, долгота, высота, пеленг, точность	Широта, долгота, высота, пеленг, точность
Wi-Fi	SSID, BSSID, тип аутентификации, IP-адрес, RSSI	SSID, уровень сигнала, информация о соединении, канал, частота	SSID, уровень сигнала, информация о соединении, канал, частота
Bluetooth	Производитель, флаг сопряжения	SSID, MAC-адрес	SSID, MAC-адрес
Датчик температуры	Значение температуры	Значение температуры	Значение температуры
Батарея	–	–	Уровень заряда
Датчик влажности	–	Уровень влажности	Уровень влажности
Микрофон	–	Аудиосигнал	–
Потребляемые ресурсы	Проценты загрузки процессора и ОЗУ	–	–
Приложение	Время запуска, продолжительность, время завершения, название приложения, флаг запуска с главного экрана	–	–

2.3. Применение цифровых отпечатков пользователя

В большинстве существующих компьютерных систем после авторизации ее ресурсы доступны пользователю до тех пор, пока он не выйдет из системы или не заблокирует сеанс. Фактически, в этот период ресурсы системы доступны любому пользователю. Например, идентификация на веб-ресурсе на основе цифрового отпечатка браузера. Однако это может привести к перехвату сеанса, когда злоумышленник нацеливается на открытый сеанс. В системах с высоким уровнем риска непрерывная идентификация личности пользователя чрезвычайно важна. Поведенческая биометрия позволяет осуществлять постоянное наблюдение за пользователями по его динамическим параметрам, например, таким как динамика нажатия [86] или динамика движения [87]. Она позволяет убедиться, что системой пользуется только авторизованный пользователь, даже после того, как была проведена первичная проверка личности. На сегодняшний день представлено множество реализаций непрерывной аутентификации от разных компаний.

Основанная в 2016 г. компания Plurilock работает над системой аутентификации с использованием поведенческой биометрии [86]. В разработанных продуктах Plurilock Aware и Plurilock Defend используется концепция жестов для аутентификации пользователей на основе динамики нажатия клавиш и движения мыши. Plurilock Aware решает проблему ввода учетных данных и избавляет от необходимости набирать пароли. Программа обеспечивает проверку личности путем распознавания шаблонов набора текста пользователей. Plurilock Defend обнаруживает легитимного пользователя, пока сессия активна, используя непрерывную аутентификацию. Если с помощью непрерывного отслеживания нажатия клавиш и мыши выявляется риск, то система получается оповещение. Эти продукты используют запатентованные алгоритмы для обеспечения непрерывной аутентификации в высокорегулируемых средах, таких как государственные учреждения, критически важные объекты инфраструктуры, финансовые услуги и здравоохранение.

Компания ThreatMark предоставляет полный пакет услуг по предотвращению текущего и будущего цифрового мошенничества [87]. ThreatMark работает над подготовкой решений для банков по борьбе с мошенничеством, начиная с раннего обнаружения угроз, поведенческой биометрии и заканчивая анализом риска транзакций.

Система идентификации голоса, разработанная компанией Aculab, захватывает десятки тысяч уникальных голосов и речевых характеристик для авторизации пользователя в режиме реального времени [88]. Утверждается, что их продукт является идеальной системой для голосовой биометрической системы аутентификации с точки зрения производительности и точности.

Система анализа поведения пользователей Cynet постоянно отслеживает и составляет профиль их активности [89]. Этот профиль впоследствии используется для определения базовой модели легитимного поведения и выявления аномальной активности, свидетельствующей о компрометации учетных записей. Система обеспечивает отслеживание всех взаимодействий в реальном времени с момента, когда пользователи начинают вход в систему.

UnifyID разработали платформу пассивной поведенческой аутентификации, предназначенной для идентификации пользователей без каких-либо осознанных действий со их стороны [90]. Разработанная платформа использует методы машинного обучения для обеспечения повышенной точности при улучшении пользовательского опыта.

SecureTouch работает над созданием технологий непрерывной аутентификации для укрепления безопасности и снижения уровня мошенничества при одновременном улучшении цифрового опыта клиентов на мобильных устройствах [91]. Их системы собирают и анализируют более чем 100 различных поведенческих параметров, таких как динамика нажатий, скорость прокрутки, давление при нажатии на экран и размер пальцев, чтобы автоматически создать уникальный поведенческий профиль пользователя, который впоследствии может быть использован для аутентификации.

Заключение

Одной из базовых задач обеспечения информационной безопасности является идентификация и аутентификация субъектов информационных процессов, в том числе протекающих в такой среде как Интернет и реализуемых с помощью коммуникационных средств Интернет-ресурсов (веб-сайтов, социальных сетей, форумов). Идентификация является основой систем разграничения доступа, и, в том числе, к Интернет-ресурсам или отдельным сервисам.

На сегодняшний день существуют две основные группы методов идентификации пользователя: по техническим характеристикам рабочей станции пользователя и по поведенческим характеристикам пользователя компьютерной системы.

Первая группа методов является хорошо проработанной и наиболее распространенной. Идентификация производится по характеристикам аппаратного и программного окружения рабочей станции пользователя, с которой осуществляется доступ ресурсу, например, через браузер. К достоинствам первой

группы методов можно отнести достаточно высокую точность идентификации, однако она обладает одним существенным недостатком: производится идентификация рабочей станции, с которой осуществляется доступ, а не конкретного пользователя этот доступ осуществляющего.

Одним из наиболее перспективных направлений развития технологий идентификации является поведенческая идентификация. Данные методы основываются на поведении пользователя. Каждый человек имеет свое уникальное поведение, которые составляют своеобразный уникальный цифровой отпечаток – набор характеристик, позволяющих его идентифицировать. Применение данного подхода позволяет произвести идентификации пользователя, а не его рабочей станции. Однако эти методы обладают несколькими недостатками и ограничениями: поведенческие характеристики пользователя обладает достаточно низкой различающей способностью.

Список источников

1. Агафонов Ю.М. Деанонимизация пользователей на основе цифровых отпечатков браузера // XVI Всероссийская научно-практическая конференция студентов, аспирантов, молодых ученых «Безопасность информационного пространства – 2017» (Екатеринбург, Российская Федерация, 12 декабря 2017). Екатеринбург: Изд-во Урал. ун-та, 2018. С. 3–5.
2. Алисултанова Э.Д., Исаева М.З., Болтиев Д.У. Анализ систем автономной идентификации пользователя сайта // Электронная наука. 2021. Т. 2. № 2. С. 7.
3. Ишкуватов С.М., Комаров И.И. Анализ аутентичности трафика на основании данных цифровых отпечатков реализаций сетевых протоколов // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 5. С. 747–754. DOI:10.17586/2226-1494-2020-20-5-747-754
4. Балмаев И.Т. Идентификация пользователей сети Tor на основе параметров перехваченного трафика // Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В. Арменского (Москва, Российская Федерация, 17 февраля – 01 марта 2017). Москва: Московский институт электроники и математики НИУ ВШЭ, 2017. С. 372–373.
5. Шулицкий Д.С., Водейко А.Э. Методы защиты авторского права на программные продукты с помощью водяных знаков и отпечатков пальцев. 2019. URL: https://libeldoc.bsuir.by/bitstream/123456789/37230/1/Shulitskiy_Metody.pdf (дата обращения 20.10.2023)
6. Гусев П.Д. Обзор существующих алгоритмов построения цифровых отпечатков // Безопасность информационных технологий. 2015. Т. 22. № 4. С. 63–67.
7. Борисова С.Н. Методы защиты аудиофайлов от несанкционированного копирования и распространения // Фундаментальные исследования. 2015. № 5-3. С. 481–487.
8. Eckersley P. How Unique Is Your Web Browser? // Proceedings of the 10th International Symposium on Privacy Enhancing Technologies Symposium (PETS 2010, Berlin, Germany, 21–23 July 2010). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2010. Vol. 6205. PP. 1–18. DOI:10.1007/978-3-642-14527-8_1
9. Mowery K., Shacham H. Pixel Perfect: Fingerprinting Canvas in HTML5 // Proceedings of W2SP. 2012. URL: https://api.semanticscholar.org/CorpusID:1399943 (Accessed 20.10.2023)
10. Cao Y., Li S., Wijmans E. (Cross-) Browser Fingerprinting via OS and Hardware Level Features // NDSS' 2017 (26 February – 1 March 2017, San Diego, USA). 2017. DOI:10.14722/ndss.2017.23152
11. Englehardt S., Narayanan A. Online tracking: A 1-million-site Measurement and Analysis // Proceedings of the Conference on Computer and Communications Security (2016 ACM SIGSAC, Vienna, Austria, 24–28 October 2016). New York: Association for Computing Machinery, 2016. PP. 1388–1401. DOI:10.1145/2976749.2978313
12. Sjösten A., Van Acker S., Sabelfeld A. Discovering Browser Extensions via Web Accessible Resources // Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy (CODASPY '17, Scottsdale, USA, 22–24 March 2017). New York: Association for Computing Machinery, 2017. PP. 329–336. DOI:10.1145/3029806.3029820
13. Sanchez-Rola I., Santos I., Balzarotti D. Extension Breakdown: Security Analysis of Browsers Extension Resources Control Policies // Proceedings of the 26th USENIX Conference on Security Symposium (USENIX Security 17, Vancouver, Canada, 16–18 August 2017). Berkeley: USENIX Association, 2017. PP. 679–694.
14. Starov O., Nikiforakis N. Xhound: Quantifying the Fingerprintability of Browser Extensions // Proceedings of the Symposium on Security and Privacy (SP, San Jose, USA, 22–26 May 2017). IEEE, 2017. PP. 941–956. DOI:10.1109/SP.2017.18
15. Mowery K., Bogenreif D., Yilek S., Shacham H. Fingerprinting Information in JavaScript Implementations. 2012. URL: https://search.iczhiku.com/paper/hgdOSDNQ7g2K8zv8.pdf (Accessed 20.10.2023)
16. Nakibly G., Shelef G., Yudilevich S. Hardware Fingerprinting Using HTML5 // arXiv:1503.01408. 2015. DOI:10.48550/arXiv.1503.01408
17. Olejnik Ł., Acar G., Castelluccia C., Diaz C. The leaking battery // Proceedings of the International Workshop on Data Privacy Management International Workshop on Quantitative Aspects in Security Assurance. 2015. С. 254–263. DOI:10.1007/978-3-319-29883-2_18
18. Vastel A., Laperdrix P., Rudametkin W., Rouvoy R. FP-STALKER: Tracking Browser Fingerprint Evolutions // Proceedings of the Symposium on Security and Privacy (SP, San Francisco, USA, 20–24 May 2018). IEEE, 2018. PP. 728–741. DOI:10.1109/SP.2018.00008
19. Laperdrix P., Rudametkin W., Baudry B. Beauty and the Beast: Diverting Modern Web Browsers to Build Unique Browser Fingerprints // Proceedings of the Symposium on Security and Privacy (SP, San Jose, USA, 22–26 May 2016). IEEE, 2016. PP. 878–894. DOI:10.1109/SP.2016.57

20. Gómez-Boix A., Laperdrix P., Baudry B. Hiding in the Crowd: an Analysis of the Effectiveness of Browser Fingerprinting at Large Scale // Proceedings of the World Wide Web Conference (Lyon, France, 23–27 April 2018). 2018. PP. 309–318. DOI:10.1145/3178876.3186097
21. A look into several Angler Exploit Kit malvertising campaigns // Operation Fingerprint. 2016. URL: https://malware bytes.app.box.com/v/operation-fingerprint (дата обращения 26.04.2023)
22. ThreatMetrix – Cybersecurity Risk Management // LexisNexis Risk Solutions. URL: https://risk.lexisnexis.com/products/ threatmetrix (дата обращения 26.04.2023)
23. The Evolution of Hi-Def Fingerprinting in Bot Mitigation // Imperva. 2018. URL: https://www.imperva.com/blog/the-evolution-of-hi-def-fingerprinting-in-bot-mitigation (дата обращения 26.04.2023)
24. Track Devices // MaxMind. 2023. URL: https://dev.maxmind.com/minfraud/track-devices (дата обращения 26.04.2023)
25. Safeguard Websites, Mobile Apps & APIs From Bots // HUMAN Bot Defender. 2023. URL: https://www.humansecurity. com/products/human-bot-defender#scroll-down (дата обращения 26.04.2023)
26. Device Fingerprint // IPQualityScore. 2023. URL: https://www.ipqualityscore.com/device-fingerprinting (дата обращения 26.04.2023)
27. Comprehensive Bot Management // Radware. 2023. URL: https://www.radware.com/products/bot-manager (дата обращения 26.04.2023)
28. Fraud Detection Software for Secure Growth // Sift. 2023. URL: https://sift.com/products/digital-trust-safety-platform (дата обращения 26.04.2023)
29. Device Recognition // SecurAuth. 2023. URL: https://docs.secureauth.com/2104/en/device-recognition.html (дата обращения 26.04.2023)
30. TruValidate // TransUnion. 2023. URL: https://www.transunion.com/solution/truvalidate (дата обращения 26.04.2023)
31. Murashko Yu.V., Papaev B.N. Increasing the accuracy of user identification by digital fingerprint using a two-stage approach // X International Scientific and Practical Conference "Prospective scientific research: experience, problems and prospects of development". 2023. С. 191-195.
32. Murashko Yu.V. Идентификация пользователей в сетевом трафике // Всероссийская научно-техническая конференция «Актуальные проблемы радиоэлектроники и телекоммуникаций» (Самара, Российская Федерация, 25–28 апреля 2023). 2023. С. 166–169.
33. Sultana M., Paul P.P., Gavrilova M. A Concept of Social Behavioral Biometrics: Motivation, Current Developments, and Future Trends // Proceedings of the International Conference on Cyberworlds (Santander, Spain, 06–08 October 2014). IEEE, 2014. PP. 271–278. DOI:10.1109/CW.2014.44
34. Мурашко Ю.В. Идентификация пользователей сети по их поведению // IX Всероссийская научно-техническая конференция «Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности» (Таганрог, Российская Федерация, 10–15 апреля 2023). Ростов-на-Дону: Южный федеральный университет, 2023. С. 54–56.
35. Rybnik M., Tabedzki M., Adamski M., Saeed K. An Exploration of Keystroke Dynamics Authentication Using Non-fixed Text of Various Length // Proceedings of the International Conference on Biometrics and Kansei Engineering (Tokyo, Japan, 05–07 July 2013). IEEE, 2013. PP. 245–250. DOI:10.1109/ICBAKE.2013.48
36. Shimshon T., Moskovitch R., Rokach L., Elovici Y. Continuous Verification Using Keystroke Dynamics // Proceedings of the International Conference on Computational Intelligence and Security (Nanning, China, 11–14 December 2010). IEEE, 2010. PP. 411–415. DOI:10.1109/CIS.2010.95
37. Teh P.S., Teoh A.B.J., Yue S. A Survey of Keystroke Dynamics Biometrics // The Scientific World Journal. 2013. Vol. 2013. P. 408280. DOI:10.1155/2013/408280
38. Tsimperidis I., Yoo P.D., Taha K., Mylonas A., Katos V. R²BN: An adaptive Model for Keystroke-Dynamics-Based Educational Level Classification // IEEE Transactions on Cybernetics. 2018. Vol. 50. Iss. 2. PP. 525–535. DOI:10.1109/TCYB.2018. 2869658
39. Alshanketi F., Traore I., Ahmed A.A. Improving Performance and Usability in Mobile Keystroke Dynamic Biometric Authentication // Proceedings of the IEEE Security and Privacy Workshops (SPW, San Jose, USA, 22–26 May 2016). IEEE, 2016. С. 66–73. DOI:10.1109/SPW.2016.12
40. Ali M.L., Thakur K., Tappert C.C., Qiu M. Keystroke Biometric User Verification Using Hidden Markov Model // Proceedings of the 3rd International Conference on Cyber Security and Cloud Computing (CSCloud, Beijing, China, 25–27 June 2016). IEEE, 2016. PP. 204–209. DOI:10.1109/CSCloud.2016.23
41. Krishnamoorthy S., Rueda L., Saad S., Elmiligi H. Identification of User Behavioral Biometrics for Authentication Using Keystroke Dynamics and Machine Learning // Proceedings of the 2nd International Conference on Biometric Engineering and Applications (ICBEA '18, Amsterdam, Netherlands, 16–18 May 2018). New York: Association for Computing Machinery, 2018. PP. 50–57. DOI:10.1145/3230820.3230829
42. Higashino J. Signature Verification System on Neuro-Computer // Proceedings of the 11th International Conference on Pattern Recognition (IAPR, Hague, Netherlands, 30 August – 3 September 1992). Vol. 3. 1992. PP. 517–521.
43. Everitt R.A.J., McOwan P.W. Java-based internet biometric authentication system // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003. Vol. 25. Iss. 9. PP. 1166–1172. DOI:10.1109/TPAMI.2003.1227991
44. Ahmed A.A.E., Traore I. A New Biometric Technology Based on Mouse Dynamics // IEEE Transactions on Dependable and Secure Computing. 2007. Vol. 4. Iss. 3. PP. 165–179. DOI:10.1109/TDSC.2007.70207
45. Antal M., Szabó L.Z. Biometric Authentication Based on Touchscreen Swipe Patterns // Procedia Technology. 2016. Vol. 22. PP. 862–869. DOI:10.1016/j.protcy.2016.01.061

46. Khan H., Hengartner U. Towards application-centric implicit authentication on smartphones // Proceedings of the 15th Workshop on Mobile Computing Systems and Applications (Santa Barbara, USA, 26–27 February 2014). New York: Association for Computing Machinery, 2014. P. 10. PP. 1–6. DOI:10.1145/2565585.2565590

47. Ferrero R., Gandino F., Montrucchio B., Rebaudengo M., Velasco A., Benkhelifa I. On gait recognition with smartphone accelerometer // Proceedings of the 4th Mediterranean Conference on Embedded Computing (MECO, Budva, Montenegro, 14–18 June 2015). IEEE, 2015. PP. 368–373. DOI:10.1109/MECO.2015.7181946

48. Fantana A.L., Ramachandran S., Schunck C.H., Talamo M. Movement based biometric authentication with smartphones // Proceedings of the International Carnahan Conference on Security Technology (ICCST, Taipei, Taiwan, 21–24 September 2015). IEEE, 2015. PP. 235–239. DOI:10.1109/CCST.2015.7389688

49. Laghari A., Waheed-ur-Rehman, Memon Z.A. Biometric authentication technique using smartphone sensor // Proceedings of the 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST, Islamabad, Pakistan, 12–16 January 2016). IEEE, 2016. PP. 381–384. DOI:10.1109/IBCAST.2016.7429906

50. Hong F., Wei M., You S., Feng Y., Guo Z. Waving authentication: your smartphone authenticate you on motion gesture // Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (Seoul, Republic of Korea, 18–23 April 2015). New York: Association for Computing Machinery, 2015. PP. 263–266. DOI:10.1145/2702613.2725444

51. Ehatisham-ul-Haq M., Azam M.A., Naeem U., Amin Y., Loo J. Continuous authentication of smartphone users based on activity pattern recognition using passive mobile sensing // Journal of Network and Computer Applications. 2018. Vol. 109. PP. 24–35. DOI:10.1016/j.jnca.2018.02.020

52. Feng T., Zhao X., Shi W. Investigating Mobile Device Picking-up motion as a novel biometric modality // Sixth International Conference on Biometrics on Theory, Applications and Systems (BTAS, Arlington, USA, 29 September – 02 October 2013). IEEE, 2013. DOI:10.1109/BTAS.2013.6712701

53. Oak R., Khare M. A Novel Architecture for Continuous Authentication Using Behavioural Biometrics // Proceedings of the International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC, Mysore, India, 8–9 September 2017). 2017. PP. 767–771. DOI:10.1109/CTCEEC.2017.8455040

54. Reynolds D.A. An Overview of Automatic Speaker Recognition Technology // Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP, Orlando, USA, 13–17 May 2002). IEEE, 2002. Vol. 4. PP. 4072–4075.

55. Ravanelli M., Bengio Y. Speaker Recognition from Raw Waveform with SincNet // Proceedings of the Spoken Language Technology Workshop (SLT, Athens, Greece, 18–21 December 2018). IEEE, 2018. PP. 1021–1028. DOI:10.1109/SLT.2018.8639585

56. Arik S., Chen J., Peng K., Ping W., Zhou Y. Neural Voice Cloning with a Few Samples // Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018, Montréal, Canada, 3–8 December 2018). 2018. Vol. 31.

57. Hu Q., Marchi E., Winarsky D., Stylianou Y., Naik D., Kajarekar S. Neural Text-to-Speech Adaptation from Low Quality Public Recordings // Proceedings of the 10th ISCA Workshop on Speech Synthesis (SSW 10, Vienna, Austria, 20–22 September 2019). 2019. Vol. 10. PP. 24–28. DOI:10.21437/SSW.2019-5

58. Kameoka H. Stargan-vc: Non-parallel many-to-many voice conversion using star generative adversarial networks // Proceedings of the Spoken Language Technology Workshop (SLT, Athens, Greece, 18–21 December 2018). IEEE, 2018. PP. 266–273. DOI:10.1109/SLT.2018.8639535

59. Rebryk Y., Beliaev S. Convoice: Real-Time Zero-Shot Voice Style Transfer with Convolutional Network // arXiv:2005.07815. 2020. DOI:10.48550/arXiv.2005.07815

60. Wu D.Y., Chen Y.H., Lee H.Y. Vqvc+: One-Shot Voice Conversion by Vector Quantization and U-Net Architecture // arXiv:2006.04154. 2020. DOI:10.48550/arXiv.2006.04154

61. Wenger E., Bronckers M., Cianfarani C., Cryan J., Sha A., Zheng H., et al. "Hello, It's Me": Deep Learning-based Speech Synthesis Attacks in the Real World // Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (Virtual Event Republic of Korea, 15–19 November 2021). New York: Association for Computing Machinery, 2021. PP. 235–251. DOI:10.1145/3460120.3484742

62. Janicki A., Alegre F., Evans N. An assessment of automatic speaker verification vulnerabilities to replay spoofing attacks // Security and Communication Networks. 2016. Vol. 9. Iss. 15. PP. 3030–3044. DOI:10.1002/sec.1499

63. Kinnunen T., Sahidullah M., Delgado H., Todisco M., Evans N. Yamagishi J., et al. The ASVspoof 2017 Challenge: Assessing the Limits of Replay Spoofing Attack Detection. 2017. URL: https://www.eurecom.fr/~evans/papers/pdfs/5235.pdf (дата обращения 22.10.2023)

64. Gao Y., Lian J., Raj B., Singh R. Detection and Evaluation of Human and Machine Generated Speech in Spoofing Attacks on Automatic Speaker Verification Systems // Proceedings of the Spoken Language Technology Workshop (SLT, Shenzhen, China, 19–22 January 2021). IEEE, 2021. PP. 544–551. DOI:10.1109/SLT48900.2021.9383558

65. Vestman V., Kinnunen T., Hautamäki R.G., Sahidullah Md. Voice Mimicry Attacks Assisted by Automatic Speaker Verification // Computer Speech & Language. 2020. Vol. 59. PP. 36–54. DOI:10.1016/j.csl.2019.05.005

66. Masuko T., Hitotsumatsu T., Tokuda K., Kobayashi T. On the security of HMM-based speaker verification systems against imposture using synthetic speech // Proceedings of the Sixth European Conference on Speech Communication and Technology (EUROSPEECH 1999, Budapest, Hungary, 5-9 September 1999). 1999.

67. Mukhopadhyay D., Shirvanian M., Saxena N. All Your Voices are Belong to Us: Stealing Voices to Fool Humans and Machines // Proceedings of the 20th European Symposium on Research in Computer Security (ESORICS 2015, Vienna, Austria, 21–25 September 2015). Lecture Notes in Computer Science. Cham: Springer, 2015. Vol. 9327. PP. 599–621. DOI:10.1007/978-3-319-24177-7_30

68. Partila P., Tovarek J., Ilk G.H., Rozhon J., Voznak M. Deep Learning Serves Voice Cloning: How Vulnerable Are Automatic Speaker Verification Systems to Spoofing Trials? // IEEE Communications Magazine. 2020. Vol. 58. Iss. 2. PP. 100–105. DOI:10.1109/MCOM.001.1900396

69. Ahmed M.E., Kwak I.Y., Huh J.H., Kim I., Oh T., Kim H. Void: A fast and light voice liveness detection system // Proceedings of the 29th USENIX Security Symposium (USENIX Security, 12–14 August 2020). 2020. PP. 2685–2702.
70. AlBadawy E.A., Lyu S., Farid H. Detecting AI-Synthesized Speech Using Bispectral Analysis // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2019. PP. 104–109.
71. Huang C., Lin Y.Y., Lee H., Lee L. et al. Defending Your Voice: Adversarial Attack on Voice Conversion // Proceedings of the Spoken Language Technology Workshop (SLT, Shenzhen, China, 19–22 January 2021). IEEE, 2021. PP. 552–559. DOI:10.1109/SLT48900.2021.9383529
72. Kobayashi R., Yamaguchi R.S. A Behavior Authentication Method Using Wi-Fi BSSIDs around Smartphone Carried by a User // Proceedings of the Third International Symposium on Computing and Networking (CANDAR, Sapporo, Japan, 08–11 December 2015). IEEE, 2015. PP. 463–469. DOI:10.1109/CANDAR.2015.45
73. Bassu D., Cochinwala M., Jain A. A new mobile biometric based upon usage context // Proceedings of the International Conference on Technologies for Homeland Security (HST, Waltham, USA, 12–14 November 2013). IEEE, 2013. PP. 441–446. DOI:10.1109/THS.2013.6699045
74. Branscomb A.S. Behaviorally Identifying Smartphone Users. Master's Thesis. Raleigh: North Carolina State University, 2013.
75. Neal T.J., Woodard D.L., Striegel A.D. Mobile device application, Bluetooth, and Wi-Fi usage data as behavioral biometric traits // Proceedings of the 7th International Conference on Biometrics Theory, Applications and Systems (BTAS, Arlington, USA, 08–11 September 2015). IEEE, 2015. PP. 1–6. DOI:10.1109/BTAS.2015.7358777
76. Fridman L., Weber S., Greenstadt R., Kam M. Active Authentication on Mobile Devices via Stylometry, Application Usage, Web Browsing, and GPS Location // IEEE Systems Journal. 2016. Vol. 11. Iss. 2. PP. 513–521. DOI:10.1109/JSYST.2015.2472579
77. Brocardo M.L., Traore I., Saad S., Woungang I. Authorship verification for short messages using stylometry // Proceedings of the International Conference on Computer, Information and Telecommunication Systems (CITS, Athens, Greece, 07–08 May 2013. IEEE, 2013. PP. 1–6. DOI:10.1109/CITS.2013.6705711
78. Cao H., Bao T., Yang Q., Chen E. An effective approach for mining mobile user habits // Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM 2010, Toronto, Canada, 26–30 October 2010). New York: Association for Computing Machinery, 2010. PP. 1677–1680. DOI:10.1145/1871437.1871702
79. Jang J., Yun J., Woo J., Kim H.K. Andro-profiler: anti-malware system based on behavior profiling of mobile malware // Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion, Seoul, Korea, 7–11 April 2014). New York: Association for Computing Machinery, 2014. PP. 737–738. DOI:10.1145/2567948.2579366
80. UMDAA-02 Dataset // Github. 2018. URL: https://umdaa02.github.io (дата обращения 26.04.2023)
81. BiDAlab/HuMIdb // Github. 2020. URL: https://github.com/BiDAlab/HuMIdb (дата обращения 26.04.2023)
82. BiDAlab/MobileB2C_BehavePassDB // Github. 2022. URL: https://github.com/BiDAlab/MobileB2C_BehavePassDB (дата обращения 26.04.2023)
83. Mahbub U., Sarkar S., Patel V.M., Chellappa R. Active user authentication for smartphones: A challenge data set and benchmark results // Proceedings of the 8th international conference on biometrics theory, applications and systems (BTAS, Niagara Falls, USA, 06–09 September 2016). IEEE, 2016. PP. 1–8. DOI:10.1109/BTAS.2016.7791155
84. Acien A., Morales A., Fierrez J., Vera-Rodriguez R., Delgado-Mohatar O. BeCAPTCHA: Behavioral bot detection using touchscreen and mobile sensors benchmarked on HuMIdb // Engineering Applications of Artificial Intelligence. 2021. Vol. 98. P. 104058. DOI:10.1016/j.engappai.2020.104058
85. Stragapede G., Vera-Rodriguez R., Tolosana R., Morales A. BehavePassDB: Public Database for Mobile Behavioral Biometrics and Benchmark Evaluation // Pattern Recognition. 2023. Vol. 134. P. 109089. DOI:10.1016/j.patcog.2022.109089
86. Ananya S.S. Keystroke Dynamics for Continuous Authentication // Proceedings of the 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence, Noida, 11–12 January 2018). IEEE, 2018. PP. 205–208. DOI:10.1109/CONFLUENCE.2018.8442703
87. Mondal S., Bours P. Continuous authentication using mouse dynamics // Proceedings of the International Conference of the BIOSIG Special Interest Group (BIOSIG, Darmstadt, Germany, 05–06 September 2013). IEEE, 2013. PP. 1–12.
88. Plurilock AI – Least privilege CASB, DLP with DEFEND technology // Plurilock. 2016. URL: https://plurilock.com (дата обращения 26.04.2023)
89. ThreatMark Fraud Prevention Solution // ThreatMark. 2015. URL: https://www.threatmark.com (дата обращения 26.04.2023)
90. VoiSentry // Aculab. 2018. URL: https://www.aculab.com/biometric-technologies/voisentry (дата обращения 26.04.2023)
91. User Behavior Analytics (UBA) // Cynet. 2018. URL: https://www.cynet.com/platform/user-behaviour-analytics (дата обращения 26.04.2023)
92. Authentication platform // UnifyID. 2015. URL: https://unify.id (дата обращения 26.04.2023)
93. SecuredTouch Company Profile – Office Locations, Competitors, Financials, Employees, Key People, News // Craft.co. 2016. URL: https://craft.co/securedtouch (дата обращения 26.04.2023)

References

1. Agafonov Yu.M. Deanonymization of users based on digital fingerprints of the browser. *Proceedings of the XVI All-Russian Scientific and Practical Conference of Students, Graduate Students, Young Scientists on Security of the Information Space – 2017, 12 December 2017, Ekaterinburg, Russian Federation.* Ekaterinburg: Ural Federal University Publ.; 2018. p.3–5.
2. Alisultanova E.D., Isaeva M.Z., Baltiev D.U. Analysis of Systemsautonomous Site User Identification. *Elektronnaya nauka.* 2021;2(2):7.

3. Ishkuvatov S.M., Komarov I.I. Traffic Authenticity Analysis Based on Digital Fingerprint Data of Network Protocol Implementations. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics.* 2020:20(5):747–754. DOI:10.17586/2226-1494-2020-20-5-747-754

4. Balmaev I.T. Identification of Tor network users based on intercepted traffic parameters. *Proceedings of the Interuniversity Scientific and Technical Conference of Students, Postgraduates and Young Specialists, 17 February – 01 March 2017, Moscow, Russian Federation*. Moscow: Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics Publ.; 2017. p.372–373.

5. Shulitsky D.S., Vodeyko A E. *Methods of Copyright Protection for Software Products Using Watermarks and Fingerprints*. 2019. URL: https://libeldoc.bsuir.by/bitstream/123456789/37230/1/Shulitskiy_Metody.pdf [Accessed 20.10.2023]

6. Gusev P.D. Review of Existing Algorithms for Constructing Digital Fingerprints. *IT Security(Russia)*. 2015;22(4):63–67.

7. Borisova S.N. Methods of Protection from the Audio-Files from Unauthorized Copying and Distribution. *Fundamental research.* 2015;3-5:481–487.

8. Eckersley P. How Unique Is Your Web Browser? *Proceedings of the 10th International Symposium on Privacy Enhancing Technologies Symposium, PETS 2010, 21–23 July 2010, Berlin, Germany. Lecture Notes in Computer Science, vol.6205*. Berlin, Heidelberg: Springer; 2010. p.1–18. DOI:10.1007/978-3-642-14527-8_1

9. Mowery K., Shacham H. Pixel Perfect: Fingerprinting Canvas in HTML5. *Proceedings of W2SP*. 2012. URL: https://api.semanticscholar.org/CorpusID:1399943 [Accessed 20.10.2023]

10. Cao Y., Li S., Wijmans E. (Cross-) Browser Fingerprinting via OS and Hardware Level Features. *NDSS' 2017, San Diego, USA, 26 February – 1 March 2017*. 2017. DOI:10.14722/ndss.2017.23152

11. Englehardt S., Narayanan A. Online tracking: A 1-million-site Measurement and Analysis. *Proceedings of the Conference on Computer and Communications Security, 2016 ACM SIGSAC, 24–28 October 2016, Vienna, Austria*. New York: Association for Computing Machinery; 2016. p.1388–1401. DOI:10.1145/2976749.2978313

12. Sjösten A., Van Acker S., Sabelfeld A. Discovering Browser Extensions via Web Accessible Resources. *Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy, CODASPY '17, 22–24 March 2017, Scottsdale, USA*. New York: Association for Computing Machinery; 2017. p.329–336. DOI:10.1145/3029806.3029820

13. Sanchez-Rola I., Santos I., Balzarotti D. Extension Breakdown: Security Analysis of Browsers Extension Resources Control Policies. *Proceedings of the 26th USENIX Conference on Security Symposium, USENIX Security 17, 16–18 August 2017, Vancouver, Canada*. Berkeley: USENIX Association; 2017. p.679–694.

14. Starov O., Nikiforakis N. Xhound: Quantifying the Fingerprintability of Browser Extensions. *Proceedings of the Symposium on Security and Privacy, SP, 22–26 May 2017, San Jose, USA)*. IEEE; 2017. p.941–956. DOI:10.1109/SP.2017.18

15. Mowery K., Bogenreif D., Yilek S., Shacham H. *Fingerprinting Information in JavaScript Implementations*. 2012. URL: https://search.iczhiku.com/paper/hgdOSDNQ7g2K8zv8.pdf [Accessed 20.10.2023]

16. Nakibly G., Shelef G., Yudilevich S. Hardware Fingerprinting Using HTML5. *arXiv:1503.01408*. 2015. DOI:10.48550/arXiv.1503.01408

17. Olejnik Ł., Acar G., Castelluccia C., Diaz C. The leaking battery. *Proceedings of the International Workshop on Data Privacy Management International Workshop on Quantitative Aspects in Security Assurance*. 2015. p.254–263. DOI:10.1007/978-3-319-29883-2_18

18. Vastel A., Laperdrix P., Rudametkin W., Rouvoy R. FP-STALKER: Tracking Browser Fingerprint Evolutions. *Proceedings of the Symposium on Security and Privacy, SP, USA, 20–24 May 2018, San Francisco*. IEEE; 2018. p.728–741. DOI:10.1109/SP.2018.00008

19. Laperdrix P., Rudametkin W., Baudry B. Beauty and the Beast: Diverting Modern Web Browsers to Build Unique Browser Fingerprints. *Proceedings of the Symposium on Security and Privacy, SP, 22–26 May 2016, San Jose, USA.* IEEE; 2016. p.878–894. DOI:10.1109/SP.2016.57

20. Gómez-Boix A., Laperdrix P., Baudry B. Hiding in the Crowd: an Analysis of the Effectiveness of Browser Fingerprinting at Large Scale. *Proceedings of the World Wide Web Conference, 23–27 April 2018, Lyon, France.* 2018. p.309–318. DOI:10.1145/3178876.3186097

21. *Operation Fingerprint*. A look into several Angler Exploit Kit malvertising campaigns. 2016. URL: https://malwarebytes.app.box.com/v/operation-fingerprint [Accessed 26.04.2023]

22. *LexisNexis Risk Solutions*. ThreatMetrix – Cybersecurity Risk Management. URL: https://risk.lexisnexis.com/products/threatmetrix [Accessed 26.04.2023]

23. *Imperva*. The Evolution of Hi-Def Fingerprinting in Bot Mitigation. 2018. URL: https://www.imperva.com/blog/the-evolution-of-hi-def-fingerprinting-in-bot-mitigation [Accessed 26.04.2023]

24. *MaxMind*. Track Devices. 2023. URL: https://dev.maxmind.com/minfraud/track-devices [Accessed 26.04.2023]

25. *HUMAN Bot Defender*. Safeguard Websites, Mobile Apps & APIs From Bots. 202]. URL: https://www.humansecurity.com/products/human-bot-defender#scroll-down [Accessed 26.04.2023]

26. *IPQualityScore*. Device Fingerprinting. 2023. URL: https://www.ipqualityscore.com/device-fingerprinting [Accessed 26.04.2023]

27. *Radware*. Comprehensive Bot Management. 2023. URL: https://www.radware.com/products/bot-manager [Accessed 26.04.2023]

28. *Sift*. Fraud Detection Software for Secure Growth. 2023. URL: https://sift.com/products/digital-trust-safety-platform [Accessed 26.04.2023]

29. *SecurAuth*. Device Recognition. 2023. URL: https://docs.secureauth.com/2104/en/device-recognition.html [Accessed 26.04.2023]

30. *TransUnion*. TruValidate. 2023. URL: https://www.transunion.com/solution/truvalidate [Accessed 26.04.2023]

31. Murashko Yu.V., Papaev B.N. Increasing the Accuracy of User Identification by Digital Fingerprint Using a Two-Stage Approach. *X International Scientific and Practical Conference on Prospective scientific research: experience, problems and prospects of development*. 2023. p.191-195.

32. Murashko Yu.V. User Identification in Network Traffic. Proceedings of the All-Russian Scientific and Technical Conference on Actual Problems of Radio Electronics and Telecommunications, 25–28 April 2023, Samara, Russian Federation. 2023. p.166-169.

33. Sultana M., Paul P.P., Gavrilova M. A Concept of Social Behavioral Biometrics: Motivation, Current Developments, and Future Trends. *Proceedings of the International Conference on Cyberworlds, 06–08 October 2014, Santander, Spain*. IEEE; 2014. p.271–278. DOI:10.1109/CW.2014.44

34. Murashko Yu.V. Identification of network users by their behavior. Proceedings of the IX All-Russian Scientific and Technical Conference on Fundamental and Applied Aspects of Computer Technologies and Information Security, 10–15 April 2023, Taganrog, Russian Federation). Rostov-on-Donu: Southern Federal University Publ.; 2023. p.54–56.

35. Rybnik M., Tabedzki M., Adamski M., Saeed K. An Exploration of Keystroke Dynamics Authentication Using Non-fixed Text of Various Length. *Proceedings of the International Conference on Biometrics and Kansei Engineering, 05–07 July 2013, Tokyo, Japan*. IEEE; 2013. p.245–250. DOI:10.1109/ICBAKE.2013.48

36. Shimshon T., Moskovitch R., Rokach L., Elovici Y. Continuous Verification Using Keystroke Dynamics. *Proceedings of the International Conference on Computational Intelligence and Security, 11–14 December 2010, Nanning, China*. IEEE; 2010. p.411–415. DOI:10.1109/CIS.2010.95

37. Teh P.S., Teoh A.B.J., Yue S. A Survey of Keystroke Dynamics Biometrics. *The Scientific World Journal*. 2013;2013:408280. DOI:10.1155/2013/408280

38. Tsimperidis I., Yoo P.D., Taha K., Mylonas A., Katos V. R²BN: An adaptive Model for Keystroke-Dynamics-Based Educational Level Classification. *IEEE Transactions on Cybernetics*. 2018;50(2):525–535. DOI:10.1109/TCYB.2018.2869658

39. Alshanketi F., Traore I., Ahmed A.A. Improving Performance and Usability in Mobile Keystroke Dynamic Biometric Authentication. *Proceedings of the IEEE Security and Privacy Workshops, SPW, 22–26 May 2016, San Jose, USA*. IEEE; 2016. p. 66–73. DOI:10.1109/SPW.2016.12

40. Ali M.L., Thakur K., Tappert C.C., Qiu M. Keystroke Biometric User Verification Using Hidden Markov Model. *Proceedings of the 3rd International Conference on Cyber Security and Cloud Computing, CSCloud, 25–27 June 2016, Beijing, China*. IEEE; 2016. p.204–209. DOI:10.1109/CSCloud.2016.23

41. Krishnamoorthy S., Rueda L., Saad S., Elmiligi H. Identification of User Behavioral Biometrics for Authentication Using Keystroke Dynamics and Machine Learning. *Proceedings of the 2nd International Conference on Biometric Engineering and Applications, ICBEA '18, 16–18 May 2018, Amsterdam, Netherlands*. New York: Association for Computing Machinery; 2018. p.50–57. DOI:10.1145/3230820.3230829

42. Higashino J. Signature Verification System on Neuro-Computer. *Proceedings of the 11th International Conference on Pattern Recognition, IAPR, 30 August – 3 September 1992, Hague, Netherlands,* vol.3. 1992. p.517–521.

43. Everitt R.A.J., McOwan P.W. Java-based internet biometric authentication system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003;25(9):1166–1172. DOI:10.1109/TPAMI.2003.1227991

44. Ahmed A.A.E., Traore I. A New Biometric Technology Based on Mouse Dynamics. *IEEE Transactions on Dependable and Secure Computing*. 2007;4(3):165–179. DOI:10.1109/TDSC.2007.70207

45. Antal M., Szabó L.Z. Biometric Authentication Based on Touchscreen Swipe Patterns. *Procedia Technology*. 2016;22: 862–869. DOI:10.1016/j.protcy.2016.01.061

46. Khan H., Hengartner U. Towards application-centric implicit authentication on smartphones. *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications, 26–27 February 2014, Santa Barbara, USA*. New York: Association for Computing Machinery; 2014. P. 10. p.1–6. DOI:10.1145/2565585.2565590

47. Ferrero R., Gandino F., Montrucchio B., Rebaudengo M., Velasco A., Benkhelifa I. On gait recognition with smartphone accelerometer. *Proceedings of the 4th Mediterranean Conference on Embedded Computing, MECO, 14–18 June 2015, Budva, Montenegro*. IEEE; 2015. p.368–373. DOI:10.1109/MECO.2015.7181946

48. Fantana A.L., Ramachandran S., Schunck C.H., Talamo M. Movement based biometric authentication with smartphones. *Proceedings of the International Carnahan Conference on Security Technology, ICCST, 21–24 September 2015, Taipei, Taiwan*. IEEE; 2015. p.235–239. DOI:10.1109/CCST.2015.7389688

49. Laghari A., Waheed-ur-Rehman, Memon Z.A. Biometric authentication technique using smartphone sensor. *Proceedings of the 13th International Bhurban Conference on Applied Sciences and Technology, IBCAST, 12–16 January 2016, Islamabad, Pakistan*. IEEE; 2016. p.381–384. DOI:10.1109/IBCAST.2016.7429906

50. Hong F., Wei M., You S., Feng Y., Guo Z. Waving authentication: your smartphone authenticate you on motion gesture. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, 18–23 April 2015, Seoul, Republic of Korea*. New York: Association for Computing Machinery; 2015. p.263–266. DOI:10.1145/2702613.2725444

51. Ehatisham-ul-Haq M., Azam,M.A., Naeem U., Amin Y., Loo J. Continuous authentication of smartphone users based on activity pattern recognition using passive mobile sensing. *Journal of Network and Computer Applications*. 2018;109:24–35. DOI:10.1016/j.jnca.2018.02.020

52. Feng T., Zhao X., Shi W. Investigating Mobile Device Picking-up motion as a novel biometric modality. *Proceedings of the Sixth International Conference on Biometrics on Theory, Applications and Systems, BTAS, 29 September – 02 October 2013, Arlington, USA.* IEEE; 2013. DOI:10.1109/BTAS.2013.6712701

53. Oak R., Khare M. A Novel Architecture for Continuous Authentication Using Behavioural Biometrics. *Proceedings of the International Conference on Current Trends in Computer, Electrical, Electronics and Communication, India, 8–9 September 2017, CTCEEC, Mysore.* 2017. p.767–771. DOI:10.1109/CTCEEC.2017.8455040

54. Reynolds D.A. An Overview of Automatic Speaker Recognition Technology. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP, 13–17 May 2002, Orlando, USA, vol.4*. IEEE; 2002. p.4072–4075.

55. Ravanelli M., Bengio Y. Speaker Recognition from Raw Waveform with SincNet. *Proceedings of the Spoken Language Technology Workshop, SLT, 18–21 December 2018, Athens, Greece*. IEEE; 2018. p.1021–1028. DOI:10.1109/SLT.2018.8639585

56. Arik S., Chen J., Peng K., Ping W., Zhou Y. Neural Voice Cloning with a Few Samples. *Proceedings of the 32nd Conference on Neural Information Processing Systems, NeurIPS 2018, 3–8 December 2018, Montréal, Canada*, vol.31. 2018.

57. Hu Q., Marchi E., Winarsky D., Stylianou Y., Naik D., Kajarekar S. Neural Text-to-Speech Adaptation from Low Quality Public Recordings. *Proceedings of the 10th ISCA Workshop on Speech Synthesis, SSW 10, 20-22 September 2019, Vienna, Austria, vol.10*. 2019. p.24–28. DOI:10.21437/SSW.2019-5

58. Kameoka H. Stargan-vc: Non-parallel many-to-many voice conversion using star generative adversarial networks. *Proceedings of the Spoken Language Technology Workshop, SLT, 18–21 December 2018, Athens, Greece*. IEEE; 2018. p.266–273. DOI:10.1109/SLT.2018.8639535

59. Rebryk Y., Beliaev S. Convoice: Real-Time Zero-Shot Voice Style Transfer with Convolutional Network. *arXiv:2005.07815*. 2020. DOI:10.48550/arXiv.2005.07815

60. Wu D.Y., Chen Y.H., Lee H.Y. Vqvc+: One-Shot Voice Conversion by Vector Quantization and U-Net Architecture. *arXiv:2006.04154*. 2020. DOI:10.48550/arXiv.2006.04154

61. Wenger E., Bronckers M., Cianfarani C., Cryan J., Sha A., Zheng H., et al. "Hello, It's Me": Deep Learning-based Speech Synthesis Attacks in the Real World. *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, 15–19 November 2021, Virtual Event Republic of Korea*. New York: Association for Computing Machinery; 2021. p.235–251. DOI:10.1145/3460120.3484742

62. Janicki A., Alegre F., Evans N. An assessment of automatic speaker verification vulnerabilities to replay spoofing attacks. *Security and Communication Networks*. 2016;9(15):3030–3044. DOI:10.1002/sec.1499

63. Kinnunen T., Sahidullah M., Delgado H., Todisco M., Evans N. Yamagishi J., et al. *The ASVspoof 2017 Challenge: Assessing the Limits of Replay Spoofing Attack Detection*. 2017. URL: https://www.eurecom.fr/~evans/papers/pdfs/5235.pdf [Accessed 22.10.2023]

64. Gao Y., Lian J., Raj B., Singh R. Detection and Evaluation of Human and Machine Generated Speech in Spoofing Attacks on Automatic Speaker Verification Systems. *Proceedings of the Spoken Language Technology Workshop, SLT, 19–22 January 2021, Shenzhen, China*. IEEE; 2021. p.544–551. DOI:10.1109/SLT48900.2021.9383558

65. Vestman V., Kinnunen T., Hautamäki R.G., Sahidullah Md. Voice Mimicry Attacks Assisted by Automatic Speaker Verification. *Computer Speech & Language*. 2020;59:36–54. DOI:10.1016/j.csl.2019.05.005

66. Masuko T., Hitotsumatsu T., Tokuda K., Kobayashi T. On the security of HMM-based speaker verification systems against imposture using synthetic speech*. Proceedings of the Sixth European Conference on Speech Communication and Technology, EUROSPEECH 1999, 5-9 September 1999, Budapest, Hungary*. 1999.

67. Mukhopadhyay D., Shirvanian M., Saxena N. All Your Voices are Belong to Us: Stealing Voices to Fool Humans and Machines. *Proceedings of the 20th European Symposium on Research in Computer Security, ESORICS 2015, 21–25 September 2015, Vienna, Austria. Lecture Notes in Computer Science,* vol.9327. Cham: Springer; 2015. p.599–621. DOI:10.1007/978-3-319-24177-7_30

68. Partila P., Tovarek J., Ilk G.H., Rozhon J., Voznak M. Deep Learning Serves Voice Cloning: How Vulnerable Are Automatic Speaker Verification Systems to Spoofing Trials? *IEEE Communications Magazine*. 2020;58(2):100–105. DOI:10.1109/MCOM.001.1900396

69. Ahmed M.E., Kwak I.Y., Huh J.H., Kim I., Oh T., Kim H. Void: A fast and light voice liveness detection system. *Proceedings of the 29th USENIX Security Symposium, USENIX Security, 12–14 August 2020.* 2020. p.2685–2702.

70. AlBadawy E.A., Lyu S., Farid H. Detecting AI-Synthesized Speech Using Bispectral Analysis. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. 2019. p.104–109.

71. Huang C., Lin Y.Y., Lee H., Lee L. et al. Defending Your Voice: Adversarial Attack on Voice Conversion. *Proceedings of the Spoken Language Technology Workshop, SLT, 19–22 January 2021, Shenzhen, China*. IEEE; 2021. p.552–559. DOI:10.1109/SLT48900.2021.9383529

72. Kobayashi R., Yamaguchi R.S. A Behavior Authentication Method Using Wi-Fi BSSIDs around Smartphone Carried by a User. *Proceedings of the Third International Symposium on Computing and Networking, CANDAR, 08–11 December 2015, Sapporo, Japan*. IEEE; 2015. p.463–469. DOI:10.1109/CANDAR.2015.45

73. Bassu D., Cochinwala M., Jain A. A new mobile biometric based upon usage context. *Proceedings of the International Conference on Technologies for Homeland Security, HST, 12–14 November 2013, Waltham, USA*. IEEE; 2013. p.441–446. DOI:10.1109/THS.2013.6699045

74. Branscomb A.S. *Behaviorally Identifying Smartphone Users*. Master's Thesis. Raleigh: North Carolina State University; 2013.

75. Neal T.J., Woodard D.L., Striegel A.D. Mobile device application, Bluetooth, and Wi-Fi usage data as behavioral biometric traits. *Proceedings of the 7th International Conference on Biometrics Theory, Applications and Systems, BTAS, 08–11 September 2015, Arlington, USA*. IEEE; 2015. p.1–6. DOI:10.1109/BTAS.2015.7358777

76. Fridman L., Weber S., Greenstadt R., Kam M. Active Authentication on Mobile Devices via Stylometry, Application Usage, Web Browsing, and GPS Location. *IEEE Systems Journal*. 2016;11(2):513–521. DOI:10.1109/JSYST.2015.2472579

77. Brocardo M.L., Traore I., Saad S., Woungang I. Authorship verification for short messages using stylometry. *Proceedings of the International Conference on Computer, Information and Telecommunication Systems, CITS, 07–08 May 2013, Athens, Greece*. IEEE; 2013. p.1–6. DOI:10.1109/CITS.2013.6705711

78. Cao H., Bao T., Yang Q., Chen E. An effective approach for mining mobile user habits. *Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM 2010, 26–30 October 2010, Toronto, Canada*. New York: Association for Computing Machinery; 2010. p.1677–1680. DOI:10.1145/1871437.1871702
79. Jang J., Yun J., Woo J., Kim H.K. Andro-profiler: anti-malware system based on behavior profiling of mobile malware. *Proceedings of the 23rd International Conference on World Wide Web, WWW '14 Companion, 7–11 April 2014, Seoul, Korea*. New York: Association for Computing Machinery, 2014. p.737–738. DOI:10.1145/2567948.2579366
80. *Github*. UMDAA-02 Dataset. 2018. URL: https://umdaa02.github.io [Accessed 26.04.2023]
81. *Github*. BiDAlab/HuMIdb. 2020. URL: https://github.com/BiDAlab/HuMIdb [Accessed 26.04.2023]
82. *Github*. BiDAlab/MobileB2C_BehavePassDB. 2022. URL: https://github.com/BiDAlab/MobileB2C_BehavePassDB [Accessed 26.04.2023]
83. Mahbub U., Sarkar S., Patel V.M., Chellappa R. Active user authentication for smartphones: A challenge data set and benchmark results. *Proceedings of the 8th international conference on biometrics theory, applications and systems, BTAS, 06–09 September 2016, Niagara Falls, USA*. IEEE; 2016. p.1–8. DOI:10.1109/BTAS.2016.7791155
84. Acien A., Morales A., Fierrez J., Vera-Rodriguez R., Delgado-Mohatar O. BeCAPTCHA: Behavioral bot detection using touchscreen and mobile sensors benchmarked on HuMIdb. *Engineering Applications of Artificial Intelligence*. 2021;98:104058. DOI:10.1016/j.engappai.2020.104058
85. Stragapede G., Vera-Rodriguez R., Tolosana R., Morales A. BehavePassDB: Public Database for Mobile Behavioral Biometrics and Benchmark Evaluation. *Pattern Recognition*. 2023;134:109089. DOI:10.1016/j.patcog.2022.109089
86. Ananya, Singh S. Keystroke Dynamics for Continuous Authentication. *Proceedings of the 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 11–12 January 2018, Noida*. IEEE; 2018. p.205–208. DOI:10.1109/CONFLUENCE.2018.8442703
87. Mondal S., Bours P. Continuous authentication using mouse dynamics. *Proceedings of the International Conference of the BIOSIG Special Interest Group, BIOSIG, 05–06 September 2013, Darmstadt, Germany*. IEEE; 2013. p.1–12.
88. *Plurilock*. Plurilock AI – Least privilege CASB, DLP with DEFEND technology. 2016. URL: https://plurilock.com [Accessed 26.04.2023]
89. *ThreatMark*. ThreatMark Fraud Prevention Solution. 2015. URL: https://www.threatmark.com [Accessed 26.04.2023]
90. *Aculab*. VoiSentry. 2018. URL: https://www.aculab.com/biometric-technologies/voisentry [Accessed 26.04.2023]
91. *Cynet*. User Behavior Analytics (UBA). 2018. URL: https://www.cynet.com/platform/user-behaviour-analytics [Accessed 26.04.2023]
92. *UnifyID*. Authentication platform. 2015. URL: https://unify.id [Accessed 26.04.2023]
93. *Craft.co*. SecuredTouch Company Profile – Office Locations, Competitors, Financials, Employees, Key People, News. 2016. URL: https://craft.co/securedtouch [Accessed 26.04.2023]

Статья поступила в редакцию 05.07.2023; одобрена после рецензирования 28.08.2023; принята к публикации 17.10.2023.
The article was submitted 05.07.2023; approved after reviewing 28.08.2023; accepted for publication 17.10.2023.

Информация об авторах:

ОСИН Андрей Владимирович
кандидат технических наук, доцент кафедры «Информационная безопасность» Московского технического университета связи и информатики
https://orcid.org/0000-0002-6384-9365

МУРАШКО Юрий Викторович
аспирант кафедры «Информационная безопасность» Московского технического университета связи и информатики
https://orcid.org/0009-0003-6448-8412

Научная статья
УДК 519.61+539.1
DOI:10.31854/1813-324X-2023-9-5-112-119

Статистические характеристики фрактальной размерности трафика IoT на примере набора данных Kitsune

Олег Иванович Шелухин, sheluhin@mail.ru
Сергей Юрьевич Рыбаков ✉, s.i.rybakov@mtuci.ru

Московский технический университет связи и информатики,
Москва, 111024, Российская Федерация

Аннотация: *В работе рассмотрен метод оценки фрактальных свойств трафика, а также проведена оценка статистических параметров фрактальной размерности (ФР) трафика IoT. Анализ реального трафика с атаками из дампа Kitsune и проведенный анализ фрактальных свойств трафика в нормальном режиме и при воздействии атак типа SSDP Flood, Mirai, OS Scan показал, что скачки ФР трафика при возникновении атак могут быть использованы при создании алгоритмов обнаружения компьютерных атак в сетях IoT. Исследования показали, что в случае онлайн-анализа сетевого трафика при оценке ФР следует отдать предпочтение модифицированному алгоритму оценки показателя Херста в скользящем окне анализа.*

Ключевые слова: показатель Херста, фрактальная размерность, трешолдинг, компьютерная атака, сетевой трафик, интернет вещей

Ссылка для цитирования: Шелухин О.И., Рыбаков С.Ю. Статистические характеристики фрактальной размерности трафика IoT на примере набора данных Kitsune // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 112–119. DOI:10.31854/1813-324X-2023-9-5-112-119

IoT Traffic Fractal Dimension Statistical Characteristics on the Kitsune Dataset Example

Oleg Shelukhin, sheluhin@mail.ru
Sergey Rybakov ✉, s.i.rybakov@mtuci.ru

Moscow Technical University of Communications and Informatics,
Moscow, 111024, Russian Federation

Abstract: *The paper considers a method for estimating the fractal properties of traffic, and also evaluates the statistical parameters of the fractal dimension of IoT traffic. An analysis of real traffic with attacks from the Kitsune dump and an analysis of the fractal properties of traffic in normal mode and under the influence of attacks such as SSDP Flood, Mirai, OS Scan showed that jumps in the fractal dimension of traffic when attacks occur can be used to create algorithms for detecting computer attacks in IoT networks. Studies have shown that in the case of online analysis of network traffic, when assessing the RF, preference should be given to the modified algorithm for estimating the Hurst exponent in a sliding analysis window.*

Keywords: Hurst exponent, fractal dimension, thresholding, computer attack, network traffic, internet of things

For citation: Shelukhin O., Rybakov S. IoT Traffic Fractal Dimension Statistical Characteristics on the Kitsune Dataset Example. *Proceedings of Telecommun. Univ.* 2023; 9(5):112–119. DOI:10.31854/1813-324X-2023-9-5-112-119

Введение

Технологии интернета вещей (IoT, *аббр. от англ.* Internet of Things) появились сравнительно недавно и за последнее десятилетие получили широкое распространение. IoT представляет собой систему взаимосвязанных компьютерных сетей и физических объектов, оборудованных множеством встроенных сенсоров, которые собирают информацию об окружающей среде. С этой целью используется специальное программное обеспечение, которое обрабатывает данные и передает информацию с датчиков по сети для последующего анализа, удаленного контроля и управления объектами IoT без участия пользователя. Также программное обеспечение обеспечивает функции хранения данных и обеспечивает доступ к ним [1, 2]. Обычно в рамках IoT присутствуют отдельные сети, каждая из которых разработана для решения конкретных задач.

Специалисты в области IoT прогнозируют ежегодное увеличение на 20 % количества «умных» устройств в период с 2020 по 2025 гг. [3]. В связи с быстрым ростом количества устройств и развитием технологии в целом появляются риски, связанные с обеспечением информационной безопасности. Устройства IoT («умные» бытовые приборы с доступом в интернет) подключены к интернету, связаны между собой и нередко становятся целью злоумышленников, желающих получить доступ к ресурсам «умных» устройств.

Поскольку устройства IoT обладают ограниченным объемом памяти и малой вычислительной мощностью, в них обычно не устанавливаются средства обеспечения безопасности от сетевых атак. Однако производители услуг и оборудования, связанных с IoT, чаще всего не придерживаются принципа сквозной информационной безопасности в основном из-за экономических факторов. Это означает, что информационной безопасности в сфере IoT обычно не уделяется должного внимания, несмотря на то, что все больше пользователей и организаций приобретают «умные» устройства: роутеры, камеры видеонаблюдения и др. К сожалению, мало кто задумывается о защите этих устройств и установке актуальных обновлений.

Существует опасность, связанная с распространением целевых кибератак на устройства IoT, и количество таких атак продолжает расти [4]. Злоумышленники, в частности, используют зараженные сети «умных» устройств для осуществления DDoS-атак или в качестве прокси-серверов для других вредоносных действий. Поэтому решения вопросов информационной безопасности должны учитываться и закладываться еще на стадии проектирования устройства или услуги, и поддерживаться на протяжении всего их жизненного цикла.

Согласно предоставленным данным [5], атаки на устройства IoT обычно не требуют сложной реализации, однако они достаточно незаметны для обычных пользователей. Одним из самых распространенных видов вредоносных программ, позволяющих ботнетам захватывать и управлять устройствами IoT путем использования устаревших уязвимостей, является Mirai. Например, одна из версий вредоносной сети Mirai проникла в более чем 5 миллионов устройств, включая устройства IoT, в 164 странах мира. Это семейство вредоносного программного обеспечения использовалось в 39 % всех атак. К числу других наиболее распространенных атак в сетях IoT относятся атаки типа SSDP Flood, OS Scan.

Постановка задачи

Одним из важных параметров сетевого трафика, который может быть положен в основу создания средства обеспечения безопасности от сетевых атак IoT, являются характеристики его фрактальных свойств. Известно, что сетевой трафик обладает свойствами самоподобия или фрактальными свойствами [6, 7]. Для количественной оценки фрактальных свойств трафика в основном используется показатель Херста H, который связан с фрактальной размерностью (ФР) D следующим соотношением: $D = 2 - H$.

В работах [8–10] было показано, что на основе показателя Херста можно обнаружить аномальную активность сетевого трафика, которая может характеризоваться следующими статистическими характеристиками:

1) выборочное среднее:

$$M_{H,i} = \frac{1}{n}\sum_{j=i}^{i+n} s_j ,$$

где s_j – выборочное значение оценки показателя Херста трафика в момент t_j;

2) выборочная дисперсия:

$$D_{H,i} = \frac{1}{n-1}\sum_{j=i}^{i+n}(s_j - M_{H,i})^2;$$

3) коэффициент асимметрии, определяющий степень асимметричности плотности вероятности распределения показателя Херста относительно оси, проходящей через центр ее тяжести:

$$\gamma_{H1,i} = \frac{\frac{1}{n-1}\sum_{j=i}^{i+n}(s_j - M_{H,i})^2}{D_{H,i}};$$

4) коэффициент эксцесса, демонстрирующий, насколько острую вершину имеет плотность распределения вероятности показателя Херста по сравнению с нормальным распределением:

$$\gamma_{H2,i} = \frac{\frac{1}{n-1}\sum_{j=i}^{i+n}(s_j - M_{H,i})^4}{D^2{}_{H,i}} - 3.$$

Данные параметры могут быть положены в основу построения эффективной системы сетевой защиты на базе интеллектуального анализа данных [11, 12] и методов фрактального анализа.

Целью работы является исследование статистических характеристик ФР наиболее распространенных атак в сети IoT на примере анализа базы данных Kitsune.

База данных

Оценку статистических параметров ФР трафика IoT будем проводить на основе выгрузки данных сетевого трафика базы Kitsune [13–15]. Kitsune – это сетевая система обнаружения вторжений (NIDS, *аббр. от англ.* Network Intrusion Detection System) на основе искусственной нейронной сети (ANN, *аббр. от англ.* Artificial Neural Network), работающей онлайн в автоматическом режиме.

На рисунке 1 представлена топология сети, на основе которой осуществлялся захват сетевого трафика на маршрутизаторах в точках 1, 2, 3 и *X*. Для каждого набора данных вначале захватывался чистый трафик для первого миллиона пакетов, а затем проводилась атака. На иллюстрации также указаны векторы атак.

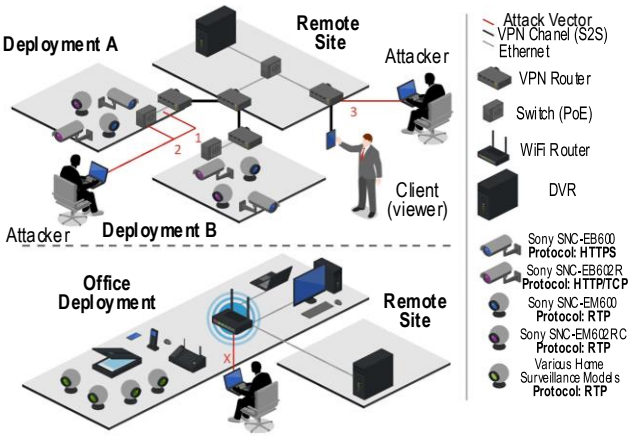


Рис. 1. Топология сети [13]

Fig. 1. Network Topology [13]

Структура извлечения объектов, называемая AfterImage, эффективно отслеживает шаблоны каждого сетевого канала, используя затухающие инкрементные статистические данные, и извлекает вектор признаков для каждого пакета. Вектор фиксирует временной контекст канала и отправителя пакета. Наблюдаемые объекты отображаются на видимые нейроны ансамбля автокодеров (KitNET https://github.com/ymirsky/KitNET-py).

Набор данных Kitsune был передан в крупнейший репозиторий реальных и модельных задач машинного обучения с протоколом UCI (*аббр. от англ.* Universal Chess Interface) и стал общедоступным в 2019 г. В нем содержится информация о четырех типах атак: разведка (Recon), человек посередине (MitM), отказ в обслуживании (DoS) и вредоносное ПО для ботнетов (Botnet Malware), например, Mirai – вредоносное программное обеспечение, которое заражает устройства IoT, работающие на процессорах ARC, и превращает их в сеть дистанционно управляемых ботов. Последних также называют «зомби». Этот ботнет часто используется для запуска DDoS-атак.

Данные об атаках были получены из коммерческой IP-системы наблюдения и сети, включающей устройства IoT. Каждый набор данных содержит миллионы сетевых пакетов и различные кибератаки.

Для каждого типа атак имеется следующий набор данных:

– предварительно обработанный набор данных, который готов для применения алгоритмов машинного обучения в формате `.csv`;

– файл с метками (также в формате `.csv`);

– исходный захваченный сетевой трафик в формате `.pcap`.

В таблице 1 представлены типы и виды сетевых атак, которые содержатся в наборе данных Kitsune. В исследовании анализировались фрактальные свойства трафика IoT до и во время воздействия атак: SSDP Flood (длительность 54 сек.), Mirai (44 мин.), OS Scan (длительность 29 сек.), представленные на рисунке 2.

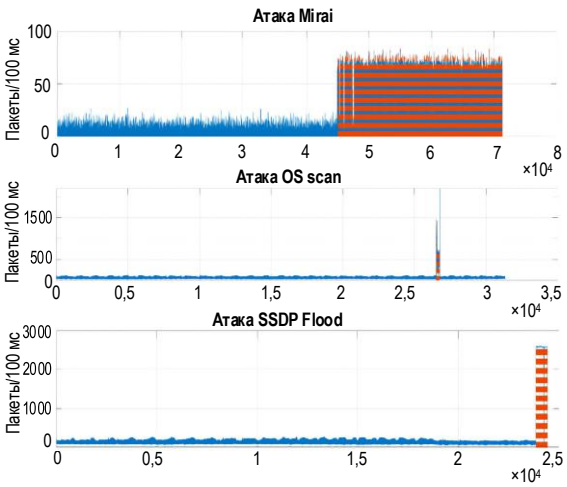


Рис. 2. Трафик IoT: нормальный (голубой); с атакой (красный)

Fig. 2. IoT Traffic: Normal Traffic (Blue), Attack Traffic (Red)

Для оценки фрактальных свойств использовался трафик пакетов, захваченный в скользящем окне с интервалом захвата в 100 мс. В таблице 2 представлены характеристики предварительно обработанного набора данных в формате `.csv`, а также вектор меток, соответствующий исходному сетевому захвату в формате `.pcap`.

Каждая строка csv-файла представляет собой перехваченный и обработанный пакет и содержит информацию о временной статистике, которая

описывает контекст передачи этого пакета, включая данные о хостах и протоколах, участвовавших в передаче. Эта информация содержит 115 различных статистических данных (атрибутов), относящихся к отправителю пакета и трафику между отправителем и получателем. Сбор статистики осуществлялся для всего трафика, который отправляется от источника, используя исходный MAC-адрес и IP-адрес пакета (SrcMAC-IP). Для получения дополнительной информации при анализе трафика использовался и исходный IP-адрес пакета (SrcIP). Для изучения канала связи между исходным и целевым IP-адресами пакета (обозначенного как канал) можно анализировать отправляемые данные. Для изучения сетевых соединений, которые обозначаются как Socket и определяются между источником и получателем пакета, использовалась информация о сокете протокола TCP/UDP.

Общее количество признаков (атрибутов), которые можно извлечь из одного временного окна анализа, составляет 23. Для извлечения атрибутов используется пять окон анализа различной длительности: 100 мс, 500 мс, 1,5 сек., 10 сек. и 1 мин., что в совокупности позволяет создать 115 атрибутов. При отсутствии данных в пакете протокола TCP/UDP соответствующие функции обнуляются.

ТАБЛИЦА 1. Информация об атаках

TABLE 1. Attack Information

Тип атаки	Название атаки	Описание	Вектор атаки	Количество пакетов	Длительность, мин.
Recon	OS Scan	Атакующий сканирует хосты в сети и их операционные системы, пытаясь обнаружить возможные уязвимости	1	1 697 851	52,2
	Fuzzing	Атакующий сканирует на наличие уязвимостей веб-сервер камер посредством отправки случайных команд	3	2 244 139	85,5
Man in the Middle	Video Injection	Злоумышленник встраивает записанное видео в общий видеопоток	1	2 472 401	33,4
	ARP MitM	Злоумышленник перехватывает весь LAN-трафик посредством ARP-атаки	1	2 504 267	28,2
	Active Wiretap	Злоумышленник перехватывает весь сетевой трафик через активную прослушку (сетевой мост), установленную на оголенном кабеле	2	4 554 925	95,6
Denial of Service	SSDP Flood	Злоумышленник перегружает видеорегистратор, заставляя камеры рассылать спам на сервер рекламными объявлениями	1	4 077 266	40,8
	SYN DoS	Злоумышленник отключает видеопоток камеры, перегружая ее веб-сервер	1	2 771 276	52,8
	SSL Renegotiation	Злоумышленник отключает видеопоток камеры, отправляя на нее множество пакетов повторного согласования SSL	1	6 084 492	65,6
Botnet Malware	Mirai	Злоумышленник заражает устройства IoT вредоносным программным обеспечением Mirai, используя учетные данные по умолчанию, а затем сканирует новую уязвимую сеть жертвы	*X*	764 137	118,9

ТАБЛИЦА 2. Характеристики набора данных Kitsune

TABLE 2. Characteristics of the Kitsune Dataset

Тип атаки	Название атаки	Количество пакетов
Вредоносное программное обеспечение для ботнетов	Mirai	764 136
Отказ в обслуживании	SSL Renegotiation	2 207 570
	SSDP Flood	4 077 265
	SYN DoS	2 771 275
Человек посередине	ARP MitM	2 504 266
	Видеоинъекция	2 472 400
	Активная прослушка	2 278 688
Разведка	Сканирование ОС	1 697 850
	Fuzzing	2 244 138

Экспериментальная оценка статистических параметров ФР

Наиболее часто для оценки показателя Херста, характеризующего ФР, используются анализ нормированного размаха (R/S-метод), график изменения дисперсии и вейвлет-анализ [6, 7].

При использовании R/S-метода для заданного набора наблюдений X со средним:

$$\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j,$$

где n – количество наблюдений, вводится понятие размаха (разности между максимальным и минимальным отклонением):

$$R(n) = \max\Delta_j - \min\Delta_j,$$

где $1 \le j \le n;\ \Delta_k = \sum_{i=1}^{k}(X_i - k\overline{X}); \forall k = \overline{1,n}$,

$$S(n) = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \overline{X}\right)^2.$$

Для многих природных явлений математическое ожидание нормированного размаха примерно равно cn^H при $n \to \infty$, где c – положительная константа, не зависящая от n. В результате показатель H можно оценить, изобразив график зависимости $\log(M\frac{R(n)}{S(n)})$ от $\log(n)$, и, используя полученные точки, подобрать по методу наименьших квадратов прямую линию с наклоном H [6, 7].

Чтобы определить количественное значение H, используется соотношение в виде:

$$H = \frac{\ln(R/S)}{\ln(n/2)}. \quad (1)$$

Для оценки ФР в режиме реального времени используется оценка показателя Херста в скользящем окне размера L. Для нейтрализации резких выбросов и уменьшения дисперсии искажений в работе [9] предлагается воспользоваться процедурой трешолдинга (*от англ.* Thresholding) – T. Под трешолдингом понимают метод очистки сигналов от шумов, основанный на вейвлет-преобразовании.

В результате использования трешолдинга формула для текущей оценки показателя Херста [9, 10] приобрела следующий вид:

$$H(t_m) = \sum_{l=1}^{L_0} a_l^{(H)}\phi_l^{(H)}(t_m) + \sum_{j=1}^{J}\sum_{l=1}^{L_j} T\left(d_{j,l}^{(H)}\right)\psi_{j,l}^{(H)}(t_m), \quad (2)$$

где $\phi_l^{(H)}(t_m)$, $\psi_{j,l}^{(H)}(t_m)$ – базисная масштабирующая и вейвлет-функция; $a_{j_0,l}^{(H)}$, $d_{j,l}^{(H)}$ – аппроксимирующие и детализирующие коэффициенты оценки показателя Херста при m-м положении окна фильтрации; $T(d_{j,l}^{(H)})$ – отфильтрованные с помощью преобразования трешолдинга детализирующие вейвлет-коэффициенты; $L_0 = 2^{J_{\max}}$, $(L_0 \le L)$; $J_{\max} = [\log_2 L]$ – максимальное число масштабов разложения; $[\log_2 L]$ – целая часть числа; масштабный коэффициент аппроксимации, равный скалярному произведению оценки показателя Херста $\hat{H}(t_m)$ и масштабной функции «самого грубого» масштаба j, смещенной на l единиц масштаба вправо от начала координат, определяется согласно выражению:

$$a_{j_0,l}^{(H)} = \left\langle \hat{H}(t_m), \phi_l^{(H)} \right\rangle;$$

вейвлет-коэффициент детализации масштаба j, равный скалярному произведению оценки показателя Херста $\hat{H}(t_m)$ и вейвлета масштаба j, смещенного на l единиц масштаба вправо от начала координат, определяется согласно выражению:

$$d_{j,l}^{(H)} = \left\langle \hat{H}(t_m), \psi_{j,l}^{(H)} \right\rangle,$$

Воспользовавшись соотношениями (1) и (2) для обработки экспериментальных данных трафика IoT, были получены статистические характеристики показателя Херста. На рисунке 3 представлены зависимости вычисленных статистических параметров ФР – M_H и D_H – для дампа нормального трафика и под атакой Mirai.

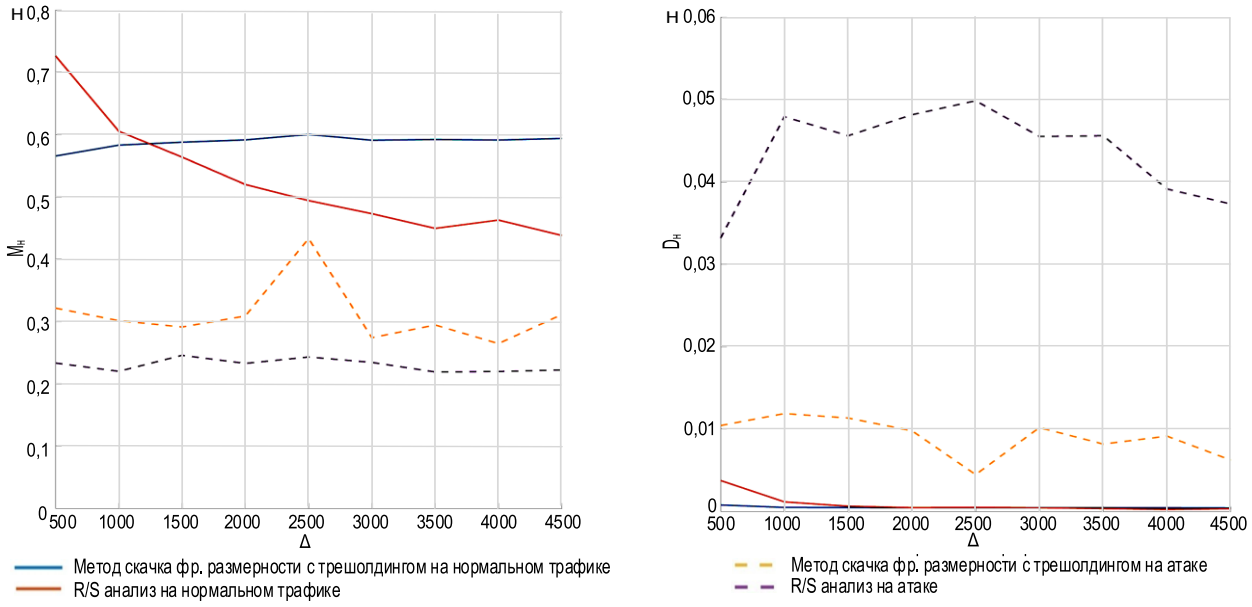


Рис. 3. Зависимости статистических параметров показателя *H* от размера окна анализа Δ для дампа нормального трафика IoT и в условиях воздействия атаки Mirai: M_H – слева; D_H – справа

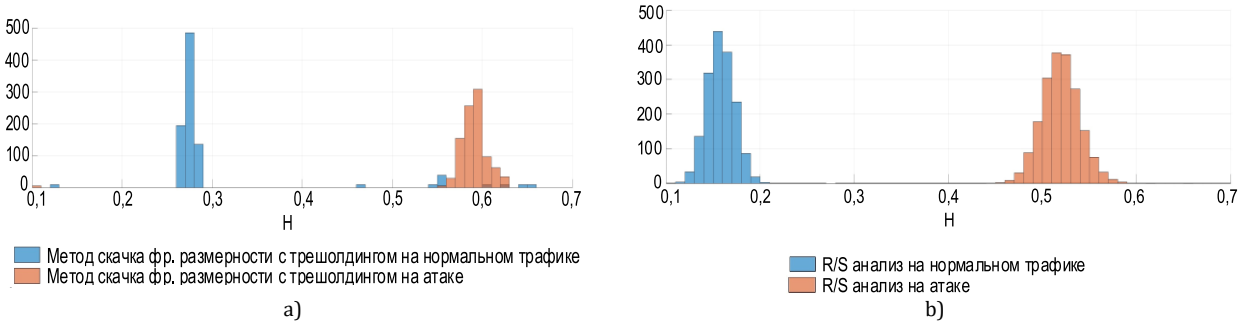
Fig. 3. Dependencies of Statistical Parameters of Indicator H on Analysis Window Size for Normal IoT Traffic Dump and under Mirai Attack: M_H – on the left; b) D_H – on the right

Сравнительный анализ представленных зависимостей показывает, что оценка указанных статистических характеристик с помощью R/S-метода и вейвлет-анализа дает в целом близкие результаты. Разброс в M_H составляет порядка 0.1, а для D_H не превышает 0.03. Различие в оценках объясняется скользящим характером оценок ФР в случае вейвлет-анализа.

Прокомментируем гистограммы распределения оценки показателя Херста на рисунке 4. Качественный анализ полученных результатов показывает, что при воздействии атаки Mirai трафик IoT имеет показатель Херста в интервале $0 < H < 0{,}5$. Это означает, что анализируемый случайный процесс не обладает самоподобием.

В свою очередь, как это видно из рисунков 3 и 4, при отсутствии атак трафик обладает фрактальными свойствами, что может быть положено в основу алгоритма обнаружения атак в сетях. На рисунке 5 показана оценка показателя Херста в скользящем окне при использовании двух рассмотренных выше алгоритмов оценки.

Атака Mirai может быть уверенно обнаружена при превышении текущей оценки показателя Херста и соответствующем выборе порогового уровня $H_{\text{пор}}$ (см. рисунки 5a и 5b). На рисунках 5c и 5d показана текущая оценка показателя Херста в скользящем окне при использовании двух алгоритмов оценки для атаки OS Scan.



a) b)

Рис. 4. Распределение показателя Херста для дампа нормального (голубой) трафика IoT и под атакой Mirai (красный) при Δ = 200 сек. и использовании метода скачка ФР с трешолдингом (a) и *R/S* анализа (b)

Fig. 4. Distribution of the Hurst Exponent for a Normal IoT Traffic Dump and a Mirai Attack at Δ = 200 sec Using the Algorithm: a) Method of Fixing Jumps of the Fractal Dimension with Thresholding; b) R/S Analysis

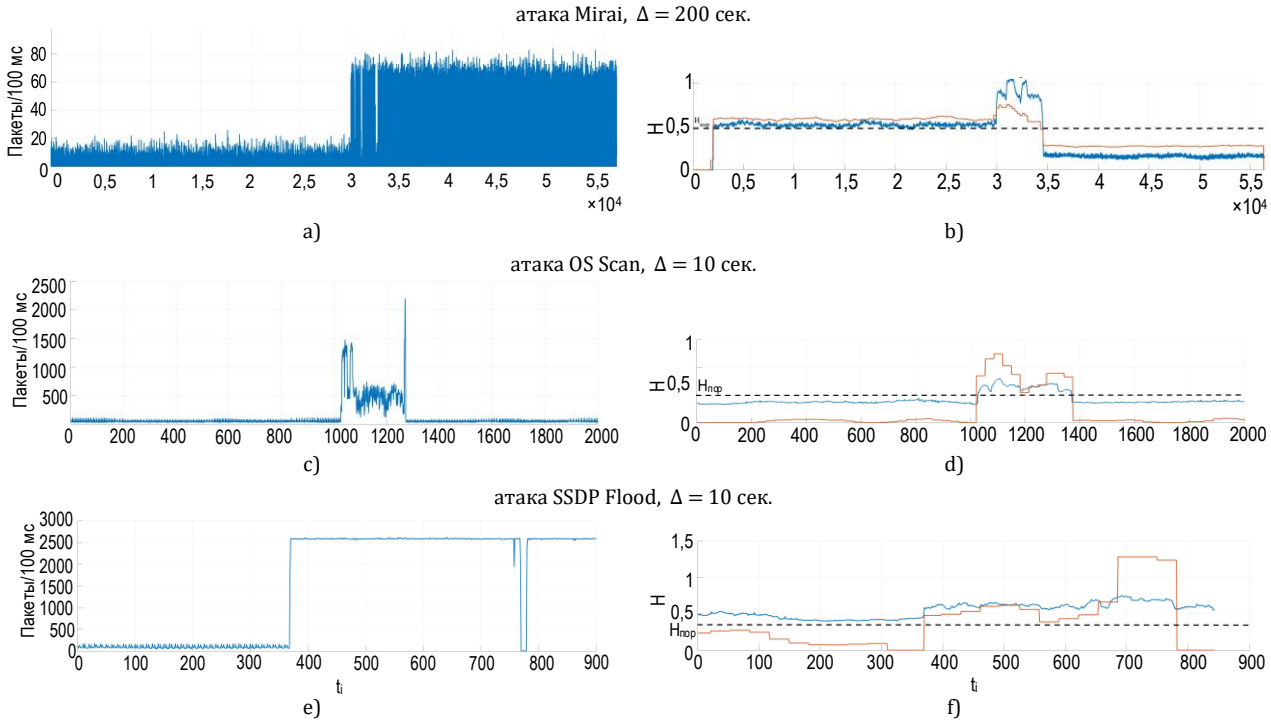


Рис. 5. Оценка *H* трафика IoT при использовании алгоритмов (1) и (2): a), c), e) фрагмент трафика с атакой; b), d), f) – оценка показателя Херста в скользящем окне

Fig. 5. Estimation H of IoT Traffic Using Algorithms (1) and (2): a), c), e) Fragment of Traffic with Attack; b), d), f) – Estimation of the Hurst Exponent in a Sliding Window

Анализ численных значений показателя Херста, представленных на рисунках 5c и 5d, показывает, что трафик IoT в отсутствие атак не обладает фрактальными свойствами, а при появлении атаки OS Scan они наблюдаются, что может быть положено в основу алгоритма обнаружения. Данное явление можно объяснить спецификой трафика устройств IoT. Как и в случае атаки Mirai, атака OS Scan может быть уверенно обнаружена при превышении текущей оценки показателя Херста порогового уровня $H_{пор}$ (см. рисунок 5d). Аналогичные результаты наблюдаются и для атаки SSDP Flood. Численное значение показателя Херста при использовании алгоритмов (1) и (2) представлены на рисунках 5e и 5f.

Сравнительный анализ зависимостей, представленных на рисунке 5, показывает, что лучшие результаты оценки ФР атак показывает алгоритм (2), реализующий метод текущей оценки ФР, основанный на вейвлет-анализе с дополнительной фильтрацией в виде преобразования трешолдинга.

Выводы

По итогам исследования были получены значения статистических параметров фрактальной размерности для нормального трафика в разных точках описанной топологии сети IoT и разных типов атак. Можно сделать вывод о том, что сетевой трафик интернета вещей обладает свойствами самоподобия в том случае, если присутствуют привычные для обычной топологии сети устройства, такие как стационарные ПК и мобильные устройства. Однако в случае, когда компьютерная сеть ограничивается лишь устройствами IoT с низкой пропускной способностью, фрактальные свойства трафика исчезают. В то же время при воздействии атак типа OS Scan и SSDP Flood у анализируемого трафика наблюдаются фрактальные свойства, что может быть использована при создании алгоритмов обнаружения компьютерных атак в сетях IoT. В случае онлайн-анализа сетевого трафика, при оценке ФР следует отдать предпочтение модифицированному алгоритму оценки показателя Херста в скользящем окне анализа (2) с использованием трешолдинга.

Список литературы

1. Minerva R., Biru A., Rotondi D. Towards a definition of the Internet of Things (IoT). Telecom Italia S.p.A., 2015. PP. 10–21. URL: https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf (Accessed 25.10.2023)

2. Dorsemaine B., Gaulier J.-P., Wary J.-P., Kheir N., Urien P. Internet of Things: A Definition & Taxonomy // Proceedings of the 9th International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST, Cambridge, UK, 09–11 September 2015). IEEE, 2015. DOI:10.1109/NGMAST.2015.71

3. Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 // Statista. URL: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide (Accessed 12.02.2023)

4. Demeter D., Preuss M., Shmelev Y. IoT: a malware story // Securelist. 2019. URL: https://securelist.com/iot-a-malware-story/94451 (Accessed 11.02.2023)

5. Шевцов В.Ю., Касимовский Н.П. Анализ угроз и уязвимостей концепций IOT и IIOT // НБИ технологии. 2020. Т. 14. № 3. С. 28–35. DOI:10.15688/NBIT.jvolsu.2020.3.5

6. Шелухин О.И. Сетевые аномалии. Обнаружение, локализация, прогнозирование. М.: Горячая линия – Телеком, 2019. 448 с.

7. Шелухин О.И. Осин А.В. Смольский С.М. Самоподобие и фракталы. Телекоммуникационные приложения. М.: Физматлит. 2008. 368 с.

8. Sheluhin O.I., Lukin I.Yu. Network traffic anomalies detection using fixing method of jumps of multifractal dimension in the real-time mode // Automatic Control and Computer Sciences. 2018. Vol. 52. Iss. 5. PP. 421–430. DOI:10.3103/S01464 11618050115

9. Шелухин О.И., Рыбаков С.Ю., Ванюшина А.В. Модификация алгоритма обнаружения сетевых атак методом фиксации скачков фрактальной размерности в режиме online // Труды учебных заведений связи. 2022. Т. 8. № 3. С. 117–126. DOI:10.31854/1813-324X-2022-8-3-117-126

10. Sheluhin O.I., Rybakov S.Y., Vanyushina A.V. Detection of Network Anomalies with the Method of Fixing Jumps of the Fractal Dimension in the Online Mode // Proceedings of the Conference on Wave Electronics and its Application in Information and Telecommunication Systems (WECONF, St. Petersburg, Russia, 30 May – 03 June 2022). IEEE, 2022. DOI:10.1109/WECONF55058.2022. 9803635

11. Sheluhin O.I., Rakovskiy D.I. Multi-Label Learning in Computer Networks // Proceedings of the Conference on Systems of Signals Generating and Processing in the Field of on Board Communications (Moscow, Russia, 14–16 March 2023). IEEE, 2023. DOI:10.1109/IEEECONF56737.2023.10092157

12. Большаков А.С., Губанкова Е.В. Обнаружение аномалий в компьютерных сетях с использованием методов машинного обучения // REDS: Телекоммуникационные устройства и системы. 2020. Т. 10. № 1. С. 37–42.

13. Mirsky Y., Doitshman T., Elovici Y., Shabtai A. Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection // arXiv:1802.09089. 2018. DOI:10.48550/arXiv.1802.09089

14. Miyamoto K., Goto H., Ishibashi R., Han C., Ban T., Takahashi, et al. Malicious Packet Classification Based on Neural Network Using Kitsune Features // Proceedings of the Second International Conference on Intelligent Systems and Pattern Recognition (ISPR 2022, Hammamet, Tunisia, 24–26 March 2022). Communications in Computer and Information Science. Cham: Springer; 2022. Vol. 1589. PP. 306–314. DOI:10.1007/978-3-031-08277-1_25

15. Alabdulatif A., Rizvi S.S.H. Machine Learning Approach for Improvement in Kitsune NID // Intelligent Automation & Soft Computing. 2022. Vol. 32. Iss. 2. PP. 827–840. DOI:10.32604/iasc.2022.021879

References

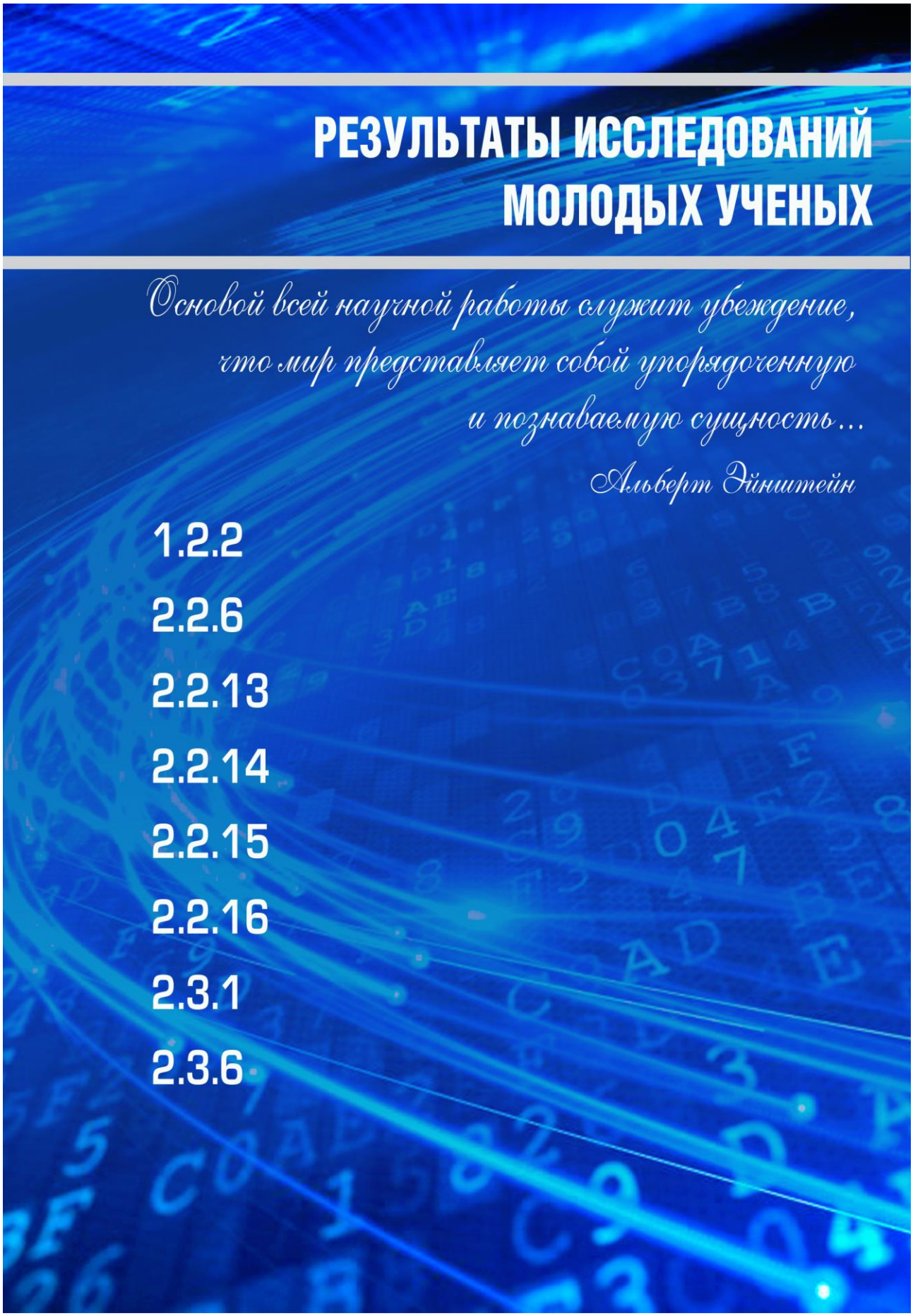
1. Minerva R., Biru A., Rotondi D. *Towards a definition of the Internet of Things (IoT).* Telecom Italia S.p.A.; 2015. p.10–21. URL: https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf [Accessed 25.10.2023]
2. Dorsemaine B., Gaulier J.-P., Wary J.-P., Kheir N., Urien P. Internet of Things: A Definition & Taxonomy. *Proceedings of the 9th International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST, Cambridge, UK, 09–11 September 2015).* IEEE; 2015. DOI:10.1109/NGMAST.2015.71
3. *Statista*. Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025. URL: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide [Accessed 12.02.2023]
4. *Securelist*. Demeter D., Preuss M., Shmelev Y. IoT: a malware story. 2019. URL: https://securelist.com/iot-a-malware-story/94451 [Accessed 11.02.2023]
5. Shevtsov V.Y., Kasimovsky N.P Threat and vulnerability analysis of IoT and IIoT concepts. *NBI technologies.* 2020;14(3): 28–35. DOI:10.15688/NBIT.jvolsu.2020.3.5
6. Sheluhin O. I. *Network Anomalies. Detection, Localization, Forecasting.* Moscow: Goryachaya liniya – Telekom Publ.; 2019. 448 p.
7. Sheluhin O.I., Osin A.V., Smolsky S.M. *Self-Similarity and Fractals*. Telecommunication. Moscow: Fizmatlit Publ.; 2008. 368 p.
8. Sheluhin O.I., Lukin I.Yu. Network traffic anomalies detection using fixing method of jumps of multifractal dimension in the real-time mode. *Automatic Control and Computer Sciences.* 2018;52(5):421–430. DOI:10.3103/S0146411618050115
9. Sheluhin O., Rybakov S., Vanyushina A. Modified Algorithm for Detecting Network Attacks Using the Fractal Dimension Jump Estimation Method in Online Mode. *Proceedings of Telecom. Univ*. 2022;8(3):117–126. DOI:10.31854/1813-324X-2022-8-3-117-126
10. Sheluhin O.I., Rybakov S.Y., Vanyushina A.V. Detection of Network Anomalies with the Method of Fixing Jumps of the Fractal Dimension in the Online Mode. *Proceedings of the Conference on Wave Electronics and its Application in Information and Telecommunication Systems. WECONF, 30 May – 03 June 2022, St. Petersburg, Russia.* IEEE; 2022. DOI:10.1109/WECONF55058.2022.9803635
11. Sheluhin O.I., Rakovskiy D.I. Multi-Label Learning in Computer Networks. *Proceedings of the Conference on Systems of Signals Generating and Processing in the Field of on Board Communications, 14–16 March 2023, Moscow, Russia.* IEEE; 2023. DOI:10.1109/IEEECONF56737.2023.10092157
12. Bolshakov A.S., Gubankova E.V. Anomaly detection in computer networks using machine learning methods. *REDS: Telecommunication Devices and Systems*. 2020;10(1):37–42.
13. Mirsky Y., Doitshman T., Elovici Y., Shabtai A. Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection. *arXiv:1802.09089*. 2018. DOI:10.48550/arXiv.1802.09089
14. Miyamoto K., Goto H., Ishibashi R., Han C., Ban T., Takahashi, et al. Malicious Packet Classification Based on Neural Network Using Kitsune Features. *Proceedings of the Second International Conference on Intelligent Systems and Pattern Recognition, ISPR 2022, 24–26 March 2022, Hammamet, Tunisia. Communications in Computer and Information Science, vol.1589*. Cham: Springer; 2022. p.306–314. DOI:10.1007/978-3-031-08277-1_25
15. Alabdulatif A., Rizvi S.S.H. Machine Learning Approach for Improvement in Kitsune NID. *Intelligent Automation & Soft Computing*. 2022;32(2):827–840. DOI:10.32604/iasc.2022.021879

Статья поступила в редакцию 18.06.2023; одобрена после рецензирования 01.08.2023; принята к публикации 02.11.2023.
The article was submitted 18.06.2023; approved after reviewing 01.08.2023; accepted for publication 02.11.2023.

Информация об авторах:

ШЕЛУХИН Олег Иванович — доктор технических наук, профессор, заведующий кафедрой «Информационная безопасность» Московского технического университета связи и информатики
https://orcid.org/0000-0001-7564-6744

РЫБАКОВ Сергей Юрьевич — аспирант кафедры «Информационная безопасность» Московского технического университета связи и информатики
https://orcid.org/0000-0002-4593-9009



DOI:10.31854/1813-324X-2023-9-5

Научная статья
УДК 004.056.52
DOI:10.31854/1813-324X-2023-9-5-121-129

Экспериментальное исследование метода защиты от атаки клонирования бумажных сертификатов

Дмитрий Алексеевич Флаксман, flxdima4951@gmail.com

ООО «Научно-производственное предприятие Новые Технологии Телекоммуникаций»,
Санкт-Петербург, 195256, Российская Федерация

Аннотация: *В работе экспериментально исследуется метод защиты бумажных сертификатов от атаки клонирования, ранее теоретически описанный в одной из работ автора. Предлагаемый метод основывается на использовании цифровых водяных знаков. Для защиты от атаки клонирования производится анализ уровня шумов, возникающих при сканировании и печати цифровых водяных знаков. В работе рассмотрены предложенные ранее алгоритмы вложения цифровых водяных знаков в изображение и последующего их извлечения, а также описывается метод выявления атаки клонирования. Приводятся результаты проведенного экспериментального вычисления вероятностей ошибок первого и второго рода для предлагаемой системы цифровых водяных знаков, которые, в основном, совпали с полученными ранее теоретическими расчетами.*

Ключевые слова: *цифровые водяные знаки, клонирование, сертификаты продукции, вероятности пропуска и ложной тревоги факта клонирования*

Благодарности: *Автор выражает благодарность профессору В.И. Коржику за постановку задачи и полезные обсуждения основных результатов работы.*

Ссылка для цитирования: Флаксман Д.А. Экспериментальное исследование метода защиты от атаки клонирования бумажных сертификатов // Труды учебных заведений связи. 2023. Т. 9. № 5. С. 121–129. DOI:10.31854/1813-324X-2023-9-5-121-129

Experimental Investigation of Protection Method for Detection of Cloning Attack on Paper Certificates

Dmitriy Flaksman, flxdima4951@gmail.com

Research and Production Enterprise “Novye Tekhnologii Telekommunikatsii”, Ltd,
St. Petersburg, 195256, Russian Federation

Abstract: *A method of paper certificate protection against a cloning attack is investigated, that was proposed recently theoretically in a paper of the same author. This method is based on the use of digital watermarks. In order to extend watermark approach to a protection against cloning attacks, it has been suggested to execute estimation of the noise power which appear during the image scanning and printing. Algorithms of embedding and extraction of watermarks out of the images are presented along with a method of detecting of the cloning attack after scanning and printing by an attacker. Numerical results of the experiments for the error probabilities of a cloning missing and a false alarm are also presented and are in agreement with theoretical results obtained before.*

Keywords: *digital watermarks, cloning, certificates, the missing and false alarm probabilities*

Acknowledgements: *The author expresses gratitude to Professor V.I. Korzhik for setting the task and useful discussions of the main results of the work.*

For citation: Flaksman D. Experimental Investigation of Protection Method for Detection of Cloning Attack on Paper Certificates. *Proc. of Telecommun. Univ.* 2023;9(5):121–129. DOI:10.31854/1813-324X-2023-9-5-121-129

1. Введение

В современном мире разнообразия рынка товаров и услуг, а также развития и доступности средств и технологий для их подделки, задача защиты авторских прав производителей является первостепенной для большинства компаний. Фальсифицированные банковские документы, счета и другие ценные бумаги – это те атаки злоумышленников, своевременное выявление которых напрямую влияет на успешную деятельность компании. При этом, во многих случаях для того, чтобы осуществлять производство поддельных образцов и документов, мошенникам достаточно иметь в своем распоряжении доступные на сегодняшний день устройства печати и сканирования. Это обстоятельство, в свою очередь, приводит к увеличению на рынке числа фальсификаций бумажных или произведенных из специального пластика сертификатов. В связи с этими обстоятельствами защита изделий, документов и различных товаров от подделок или ложных утверждений уникальных качеств, несомненно, является важной составляющей наиболее актуальных задач в сфере информационной безопасности.

На сегодняшний день распространенным методом борьбы с подобными проблемами является метод бумажных (или пластиковых) сертификатов, подразумевающий использование штрих-кода (например, QR-код или DataMatrix). Тем не менее, указанный метод работает не во всех случаях. Например, если производится «клонирование» сертификата (т. е. сканирование или фотографирование) и затем на основе такого клона создается поддельный, то использование такого сертификата вместе с товаром пониженного качества может остаться не обнаруженным. При этом кажущаяся визуальная подлинность сертификата гарантирует его реализацию по цене оригинала.

С целью повышения надежности сертификатов возможно также использование метода вложения цифровых водяных знаков (ЦВЗ). В этом случае, для установки подлинности будет необходим дополнительный конфиденциальный цифровой ключ, доступный исключительно собственнику продукта [1]. Но, несмотря на свои преимущества, указанный метод не гарантирует защиты изделий от возможного «клонирования» сертификатов.

Еще одним подходом в использовании вложений ЦВЗ для решения указанной проблемы является искажение отдельных блоков информации, включая штрих-коды. Данный метод описан в работе [2]. Другой подход предполагает изменение фона на специальный текстурный шаблон на основе гауссовского шума, характеристики которого чувствительны к процедуре печати и сканирования [3]. Однако стоит учитывать, что подобные искажения существенно влияют на структуру обрабатываемого объекта.

Метод, предложенный в статье [4], использует анализ различных особенностей изображения в частотной и пространственной областях. Для этого в первом случае используется коррекция изображения, определение набора локальных максимумов для оригинала и его последующее сравнение с набором аналогичных характеристик подделки. Во втором случае, к оригиналу и подделке применяется преобразование для получения изображений определенного вида, в результате которого оценивается равномерность цвета, яркость и расстояния между их частями. Так же для выявления фальсифицированной продукции проводятся исследования применения методов машинного обучения [5]. Однако для эффективной работы нейронных сетей требуется подготовка большого объема исходных обучающих данных.

В настоящей работе исследуется оригинальный метод защиты бумажных сертификатов, который позволяет выявить попытку «клонирования». Рассматриваемый метод основывается на использовании ЦВЗ и на том факте, что любые дополнительные операции над изображением, будь то сканирование, печать или цифровая обработка изображения, неминуемо приведут к увеличению мощности шума в подделываемом злоумышленником сертификате.

Предлагаемый метод подходит как для цветных изображений, так и для изображений в градациях серого, в том числе и для матричных штриховых кодов, таких как QR-код или DataMatrix.

Дальнейшие результаты структурированы следующим образом:

– во втором разделе представлено общее описание работы системы ЦВЗ;

– в третьем разделе рассматривается алгоритм вложения ЦВЗ, использующий в своей основе дискретно-косинусное преобразование и широкополосные сигналы, как это предполагалось ранее в работе [6, 7];

– в четвертом разделе описывается метод устранения геометрических искажений и алгоритм извлечения вложенных данных;

– в пятом разделе рассматривается алгоритм обнаружения атаки клонирования, основанный на оценке шумов изображения;

– в шестом разделе представлены результаты экспериментов, подтверждающих работоспособность предлагаемого метода;

– в заключении подводятся итоги полученных результатов и предлагаются возможные направления для проведения дальнейшего исследования.

2. Общая схема вложения ЦВЗ и метода клонирования

Рассмотрим общую схему работы предлагаемой системы ЦВЗ (рисунок 1).

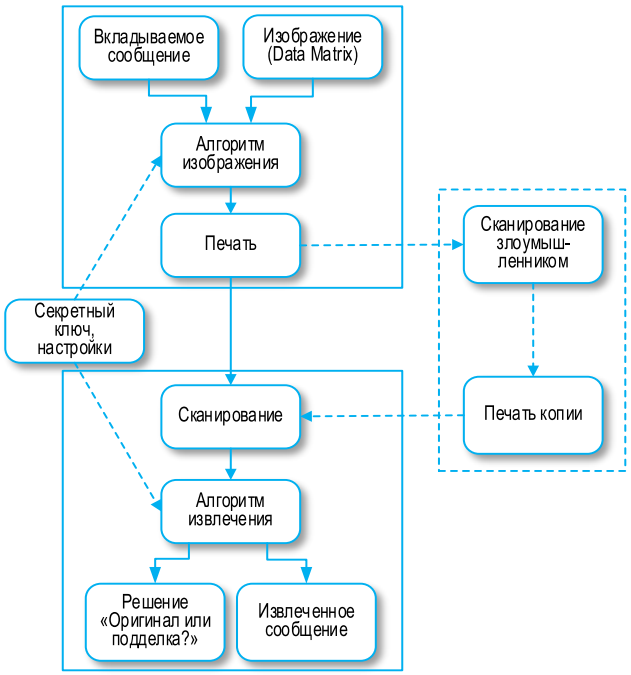


Рис. 1. Общая схема использования системы ЦВЗ и процедуры клонирования

Fig. 1. General Scheme of Using the DW System and Cloning Procedure

На вход алгоритма вложения поступает вкладываемое сообщение, покрывающее изображение, секретный ключ, а также различные второстепенные настройки алгоритма. Размер вкладываемого сообщения зависит от размеров будущей стеганограммы и физического носителя изображения. Например, для изображения 460×460 точек при печати на бумагу в размере 4×4 см можно вложить 128 бит. На выходе алгоритма вложения получается защищенное изображение (стеганограмма), которое распечатывается на физический носитель (бумагу, пластик и т. п.). Распечатанную стеганограмму назовем сертификатом.

Сертификат поступает на вход алгоритма извлечения, однако на вход также может прийти и поддельный сертификат (копия оригинального). Для работы алгоритма извлечения также необходимо знать секретный ключ. Результатом работы алгоритма извлечения является вложенное сообщение, а также решение о наличии факта подделки.

3. Алгоритм вложения

Разберем подробнее алгоритм вложения. Общая схема алгоритма представлена на рисунке 2.

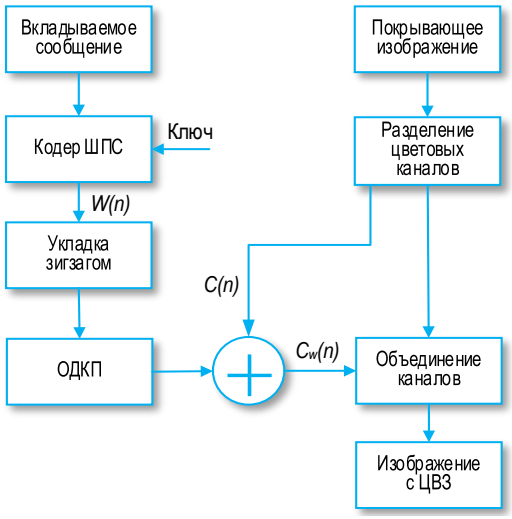


Рис. 2. Блок-схема алгоритма вложения ЦВЗ

Fig. 2. Scheme of DW Embedding

Покрывающим изображением может выступать как цветное, так и изображение в градациях серого. В частности, в этом качестве может выступать двумерный матричный штрихкод, например, DataMatrix (см. рисунок 3). Если защищается цветное изображение, то для вложения будет использоваться только его синий цветовой канал. Покрывающее сообщение обозначим как $C(n)$.

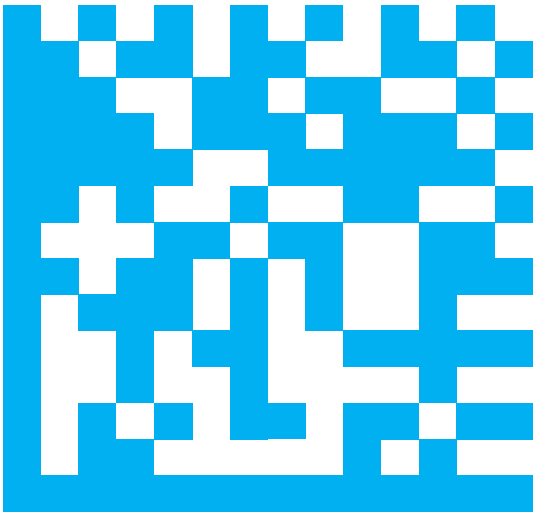


Рис. 3. Изображение DataMatrix кода

Fig. 3. Example of DataMatrix Barcode

Вкладываемое сообщение кодируется с использованием кодера широкополосных сигналов (ШПС). В кодере ШПС используется псевдослучайный сигнал, вырабатываемый на основе секретного ключа с помощью линейного рекуррентного регистра (ЛРР):

$$W^{b_i}(n) = \alpha(-1)^{b_i}\pi(n), \qquad n = 1, 2, \dots, N_0, \qquad (1)$$

где α – глубина вложения; N_0 – длина псевдослучайной последовательности (ПСП), на которой вкладывается один бит информации; π(n) – отсчет ПСП (±1); $b \in (0,1)$ – бит вкладываемого сообщения с индексом i.

Далее полученная последовательность укладывается зигзагом с таким расчетом, чтобы заполнить область «средних частот» дискретно-косинусного преобразования (ДКП). Выбор частотной области ДКП обусловлен тем, что вложение в нее будет наиболее устойчиво к искажениям. К полученной матрице применяется обратное дискретно-косинусное преобразование (ОДКП), в результате которого получается матрица $Cw(n)$. Для получения стеганограммы матрица $Cw(n)$ складывается с покрывающим изображением $C(n)$.

4. Алгоритм извлечения

Блок–схема алгоритма извлечения ЦВЗ представлена на рисунке 4. На вход алгоритма извлечения поступает отсканированное изображение.

Рис. 4. Блок-схема алгоритма извлечения ЦВЗ

Fig. 4. Scheme of DW Extraction

Для надежного извлечения вложенных данных необходимо максимально точно восстановить размер и ориентацию изображения.

В целях устранения искажений используется перспективное преобразование изображений [8]:

$$\begin{pmatrix} \bar{x}' \\ \bar{y}' \\ w \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \\ 1 \end{pmatrix}, \quad (2)$$

где x', y' – новые координаты точки; x, y – старые координаты точки; $t^{i,j}$ – коэффициенты матрицы преобразования; w – глубина (масштаб).

Для того, чтобы получить коэффициенты преобразования, можно воспользоваться различными методами. Если изображение представляет собой матричный штрихкод, то можно ориентироваться на его структуру, которая предназначена для автоматического позиционирования. При наличии оригинала для сопоставления положения можно воспользоваться методами компьютерного зрения. В крайнем случае, положение можно выбрать визуально в ручном режиме. Подробнее устранение геометрических искажений рассматривалось в статье [8].

После применения перспективного преобразования, для извлечения вложенных данных необходимо вернуться в частотную область. Далее из области средних частот ДКП зигзагом выбирается последовательность отсчетов, которая отправляется на декодер ШПС. Если есть оригинальное изображение, то можно воспользоваться *информированным* декодером ШПС или же *слепым*, если оригинал отсутствует.

5. Алгоритм обнаружения факта клонирования сертификата

Само по себе верное извлечение ЦВЗ не гарантирует того, что мы извлекаем данные из оригинального сертификата, так как он мог быть скопирован и повторно напечатан. Однако любая дополнительная операция, которая при этом производится над ЦВЗ, неминуемо влечет за собой увеличение уровня шумов, в то время как попытка удаления шумов сканирования с высокой вероятностью приведет к повреждению вложения и невозможности его извлечения.

Значение отсчета проверяемой ЦВЗ, в случае, если не было клонирования, имеет вид:

$$C_t'(n) = C_w(n) + N_{p1}(n) + N_{s1}(n), n = 1, 2..N, \quad (3)$$

где $C_w(n)$ – отчеты оригинального ЦВЗ; $N_{p1}(n)$ – шумы печати; $N_{s1}(n)$ – шумы сканирования; N – длина тестируемой последовательности.

Если клонирование произошло, то необратимо появляются дополнительные шумы:

$$C_t''(n) = C_w(n) + N_{p1}(n) + N_{s2}(n) + \\ + N_{p2}(n) + N_{s1}'(n),\ n = 1, 2..N, \quad (4)$$

где $N_{s2}(n)$ и $N_{p2}(n)$ – шумы сканирования и печати злоумышленником; $N_{s1}'(n)$ – шумы сканирования поддельного сертификата.

Для проверки подлинности легитимный пользователь рассчитывает величину:

$$\lambda(n) = C_t(n) - C_w(n), n = 1, 2..N, \quad (5)$$

где $C_t(n)$ – отсчеты проверяемого ЦВЗ.

После этого он рассчитывает нормированную мощность шумов:

$$\Omega = \frac{1}{N}\sum_{n=1}^{N} \lambda^2(n), \quad (6)$$

где N – длина тестируемой области.

При этом, зная параметры оборудования, на котором производился оригинальный сертификат, можно попытаться подобрать порог для мощности шумов, тогда решение о наличии клонирования будет приниматься по правилу:

$$\begin{cases} \Omega \geq \Omega_0 \Rightarrow \text{клонирование есть} \\ \Omega < \Omega_0 \Rightarrow \text{клонирования нет} \end{cases}, \quad (7)$$

где Ω_0 – некоторый заранее заданный порог.

При принятии решения могут появиться два вида ошибок: P_m – вероятность пропуска клонирования, когда оно в действительности произошло, но не было обнаружено; P_{fa} – вероятность ложной тревоги, когда клонирования не было, но по правилу принимается решение о его наличии.

Определим полную вероятность ошибки как

$$P_e = \frac{1}{2}(P_m + P_{fa})$$

и будем называть оптимальным порогом такую величину $\Omega = \Omega_0$, которая обеспечивает минимум P_e.

Подробный вывод выражения для вероятности ошибок представлен в статье [6]:

$$P_e \approx \Phi\left(\frac{1}{2\pi}\frac{e^{-\frac{x^2}{2}}}{x}\right), \quad x = \frac{\sqrt{N}}{\sqrt{2}(2r+1)}, \quad (8)$$

где Ф – функция Лапласа; r – это величина, которая показывает, во сколько раз дисперсия шумов у атакующего меньше дисперсии шумов у легального пользователя.

Результаты расчета P_e по (8) для различных значений r и N представлены в таблице 1 и на рисунке 5.

ТАБЛИЦА 1. Теоретическая вероятность ошибки при различных значениях *r* и *N* (%)

TABLE 1. Theoretical Error Probability for Different Values of r and N (%)

r \ N	600	800	1000	1200	1500
1	0	0	0	0	0
2	0,0046	0,00053	0,000065	0,0000079	0,00000036
4	0,52	0,24	0,12	0,058	0,021
8	3,7	2,7	2,1	1,6	1,1
16	11	8,7	7,4	6,5	5,4

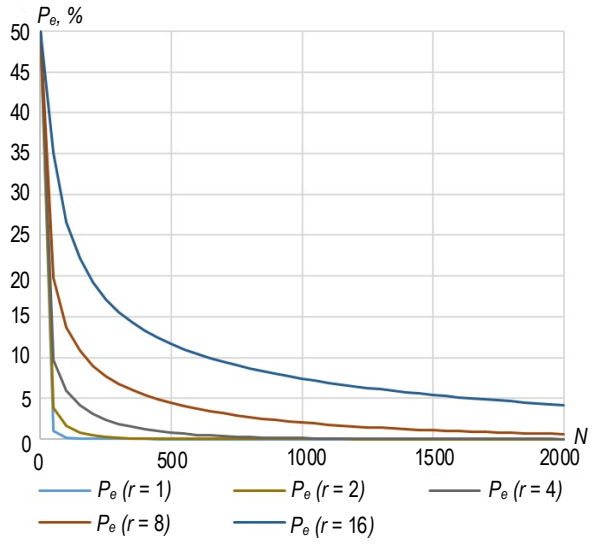


Рис. 5. Зависимость теоретической вероятности ошибки от длины тестируемой области *N*

Fig 5. Dependency of the Theoretical Error Probability against N

6. Экспериментальные результаты

Перейдем к тестированию системы ЦВЗ. Для начала проверим качество извлечения ЦВЗ. Для теста был сгенерирован двумерный матричный штрихкод типа DataMatrix с размером модуля 22×22. В нем закодировано текстовое сообщение из 32 знаков (размер изображения на основе кода – 460×460 точек):

DataMatrix*x*DataMatrix*x*DataMatrix32

В полученное изображение была произведена серия вложений ЦВЗ с различными параметрами глубины вложения α (α = 4, 7, 10) и длины ШПС (N = 300, 500, 700, 900).

На первом этапе эксперимента каждый сертификат был напечатан 24 раза на одном листе полуглянцевой бумаги формата А4. Для печати использовался струйный шестицветный фотопринтер Epson L805. Для сканирования использовался сканер Canon Lido 220. Физический размер изображения на бумаге 4×4 см.

После этого было проведено сканирование напечатанных изображений. Для наглядности было выбрано различное разрешение сканирования (200, 300, 400, 600 и 1200 dpi). Результаты извлечения 64, 128 и 256 бит вложенных данных представлены в таблице 2: для наглядности жирным шрифтом выделены результаты экспериментов, в которых средний процент ошибочно извлеченных бит менее 1 %.

По представленным данным можно увидеть, что если требуется вложить 64 бита данных, то можно использовать практически любую совокупность параметров из заданного диапазона, получая приемлемое качество извлечения, при этом разреше-

ние сканирования не оказывает значительного влияния на результат. Это обусловлено тем, что все вложение попадает в оптимальную для него область средних частот ДКП.

При использовании 128 бит данных набор подходящих параметров для вложения значительно сужается. Так, глубина вложения, равная 3, уже не кажется подходящей для использования. Так же при разрешении сканирования в 200 dpi наблюдается рост ошибочно извлеченных данных. Это косвенно говорит о том, что при увеличении части спектра ДКП, используемой для вложения, начинает захватываться область верхних частот, которая, в свою очередь, более чувствительна к искажениям.

При попытке использования 256 бит можно заметить, что алгоритм становится наиболее чувствительным к качеству отсканированного изображения, так как само вложение осуществляется в том числе и в высокие частоты ДКП. В то же время, при значении длины ШПС N = 900 происходит превышение максимальной емкости вложения.

На втором этапе эксперимента была имитирована работа злоумышленника. Для этого из отсканированных с разрешением 1200 dpi изображений были подготовлены поддельные сертификаты. При этом для того, чтобы клонированное изображение визуально не отличалось от оригинала, в фоторедакторе была произведена незначительная правка гистограммы распределения цветов. Гистограмма была подогнана под значения оригинальной стеганограммы, которой в реальности у злоумышленника в распоряжении не будет. Каждый поддельный сертификат был распечатан с такими же условиями печати и на такой же бумаге, как и на первом этапе эксперимента.

ТАБЛИЦА 2. Средний процент ошибок при вложении 64/128/256 бит данных

TABLE 2. Average Error Percent in the Case of 64/128/256 Bits Embedding

α	N	Доля ошибочно извлеченных бит, %				
		dpi – разрешение сканирования				
		200	300	400	600	1200
4	300	7,81 / 6,12 / 14,65	7,16 / 3,88 / 4,52	5,54 / 2,83 / 2,24	7,36 / 3,81 / 2,99	7,61 / 3,98 / 2,49
	500	3,19 / 5,11 / 2,73	3,00 / 1,76 / 5,84	3,33 / 1,69 / 3,76	3,13 / 1,56 / 1,07	3,39 / 1,69 / 1,33
	700	1,37 / 10,42 / 28,50	**0,52** / 1,46 / 8,01	**0,33** / 1,11 / 8,28	**0,59** / **0,29** / 3,34	**0,85** / **0,42** / 1,56
	900	1,76 / 18,46 / –	**0,19** / 1,30 / –	**0,19** / 1,01 / –	**0,33** / **0,46** / –	**0,54** / **0,41** / –
7	300	2,47 / 2,90 / 9,65	2,21 / 1,10 / 1,02	2,35 / 1,17 / 1,16	2,35 / 1,17 / **0,58**	3,06 / 1,53 / **0,76**
	500	1,30 / 1,82 / 18,47	1,76 / **0,88** / 1,17	1,70 / **0,85** / 1,92	2,08 / 1,04 / **0,55**	2,45 / 1,22 / **0,63**
	700	**0,00** / 3,29 / 22,07	**0,00** / **0,10** / 4,07	**0,00** / **0,07** / 2,90	**0,00** / **0,00** / **0,70**	**0,00** / **0,00** / 1,03
	900	1,63 / 12,50 / –	**0,00** / **0,62** / –	**0,00** / **0,00** / –	**0,00** / **0,00** / –	**0,00** / **0,00** / –
10	300	1,56 / **0,84** / 3,34	1,63 / **0,81** / **0,60**	1,63 / **0,81** / **0,80**	1,63 / **0,81** / **0,41**	1,56 / **0,78** / **0,39**
	500	**0,39** / 3,78 / 19,47	**0,39** / **0,20** / 1,82	**0,52** / **0,26** / 1,04	**0,72** / **0,36** / **0,18**	**1,17** / **0,58** / **0,32**
	700	**0,00** / 3,26 / 21,83	**0,00** / **0,00** / **3,60**	**0,00** / **0,23** / 4,17	**0,00** / **0,00** / **0,24**	**0,00** / **0,00** / **0,59**
	900	**0,00** / 7,13 / –	**0,00** / **0,00** / –	**0,00** / **0,16** / –	**0,00** / **0,75** / –	**0,00** / **0,00** / –

Поддельные сертификаты были так же отсканированы с различным разрешением сканирования. Результаты извлечения вложенных данных представлены в таблице 3: для наглядности жирным шрифтом выделены результаты экспериментов, в которых средний процент ошибочно извлеченных бит менее 2 %.

По полученным результатам можно увидеть, что количество ошибочно извлеченных бит увеличилось. Так, при вложении 256 бит данных на успешное извлечение можно рассчитывать только при определенных наборах параметров. Однако в случае вложения 64 и 128 бит количество ошибок растет не так значительно, поэтому злоумышленник может рассчитывать, что подделка будет успешной. На практике же, небольшой процент ошибочных бит будет исправлен кодом с исправлением ошибок, и подделка успешно пройдет процедуру извлечения.

Для выявления подделки перейдем к измерению шумов. Статистика была получена по результатам представленных выше экспериментов по печати и последующему клонированию сертификата. Измерение шума производилось в частотном представлении только в области средних частот ДКП, примерно соответствующих области вложения.

Для наглядности результаты расчетов по выбору оптимального порога представлены на графиках зависимости вероятностей ошибок P_m и P_{fa} от значения порога Ω_0. Так как дисперсия шума зависит от разрешения сканирования, то подбор порога удобно осуществлять для каждого разрешения сканирования по отдельности.

ТАБЛИЦА 3. Средний процент ошибок при вложении 64/128/256 бит данных при атаке клонирования

TABLE 3. Average Error Percent in the Case of 64/128/256 Bits Embedding in Case of a Cloning Attack

α	*N*	Доля ошибочно извлеченных бит, %				
		dpi – разрешение сканирования				
		200	300	400	600	1200
4	300	14,19 / 18,85 / 28,37	11,85 / 12,11 / 18,06	9,31 / 10,81 / 17,88	9,51 / 8,76 / 14,06	9,38 / 8,29 / 13,14
	500	12,57 / 23,76 / 34,67	5,54 / 11,39 / 23,25	4,69 / 10,06 / 21,26	4,17 / 6,06 / 16,20	3,78 / 6,09 / 15,36
	700	10,94 / 26,33 / 36,42	4,17 / 15,95 / 28,39	2,86 / 12,83 / 25,74	**1,76** / 9,77 / 22,64	**1,56** / 7,98 / 21,17
	900	13,48 / 29,04 / –	5,47 / 18,43 / –	4,75 / 18,17 / –	**1,95** / 13,25 / –	2,51 / 13,35 / –
7	300	3,19 / 4,88 / 15,77	3,39 / 2,51 / 5,88	2,74 / 2,70 / 8,41	3,06 / 2,11 / 5,43	2,86 / **1,80** / 3,55
	500	2,67 / 10,19 / 25,30	2,28 / 2,90 / 13,19	**1,83** / 3,97 / 12,44	2,41 / 2,96 / 9,97	2,11 / **1,83** / 7,34
	700	4,82 / 20,02 / 31,96	**0,33** / 5,44 / 18,66	**0,65** / 5,73 / 18,68	**0,07** / 3,78 / 17,78	**0,07** / 2,86 / 14,27
	900	7,75 / 25,16 / –	**1,37** / 10,81 / –	**1,43** / 11,49 / –	**0,07** / 5,26 / –	**0,14** / 5,97 / –
10	300	**1,95** / 2,64 / 11,77	**1,95** / **1,11** / 2,86	**1,76** / **1,11** / 3,81	**1,69** / **0,98** / **1,84**	**1,63** / **0,81** / **0,97**
	500	2,15 / 9,24 / 23,59	**0,65** / 5,73 / 16,15	**1,11** / 5,70 / 15,36	**0,72** / **1,76** / 7,42	**0,78** / **0,75** / 6,19
	700	2,60 / 17,42 / 29,84	**0,59** / 4,39 / 18,27	**0,07** / 5,08 / 17,15	**0,00** / **1,56** / 12,61	**0,00** / **1,21** / 10,45
	900	5,47 / 22,63 / –	**0,26** / 8,58 / –	**0,65** / 8,79 / –	**1,17** / 9,70 / –	**0,21** / 5,61 / –

Графики для значений разрешения 200, 300 и 600 dpi представлены на рисунке 6, соответственно. Статистика для каждого графика посчитана по результатам 288 экспериментов. По графику для разрешения сканирования в 200 dpi можно сделать вывод, что подобрать порог Ω_0, при котором общая вероятность ошибки будет близка к нулю, невозможно. Так, при значении порога Ω_0 = 1270 общая вероятность $P_e \approx 19$ %. Однако при разрешении сканирования в 300 dpi уже можно выбрать порог $\Omega_0 \approx 1050$, при котором общая вероятность ошибки становится P_e = 4 %. Для разрешения сканирования в 600 dpi можно подобрать порог, при котором общая вероятность ошибки будет близка к нулю. Например, при значении порога Ω_0 = 950 общая вероятность $P_e \approx 0$ %.

Из графиков (см. рисунок 6) можно заключить, что при увеличении разрешения сканирования становится проще отделить оригинальные стеганограммы от поддельных. Для наглядности построим график зависимости дисперсии шума от разрешения сканирования (рисунок 7), который показывает, что при малых значениях разрешения сканирования графики сближаются. Это обстоятельство также подтверждает, что при достаточно высоком качестве оборудования у проверяющего (или, как минимум, сопоставимым с оборудованием злоумышленника), появляется надежный метод обнаружения поддельных сертификатов, так как попытка клонировать сертификат неминуемо внесет дополнительный шум, уровень которого не удастся замаскировать.

Расчет дисперсии и порога Ω_0 в представленных выше экспериментах производился для области средних частот, при этом длина тестируемой области *N* составляла 31740 отсчетов. В теоретических же расчетах значение *N* не превышало 2000.

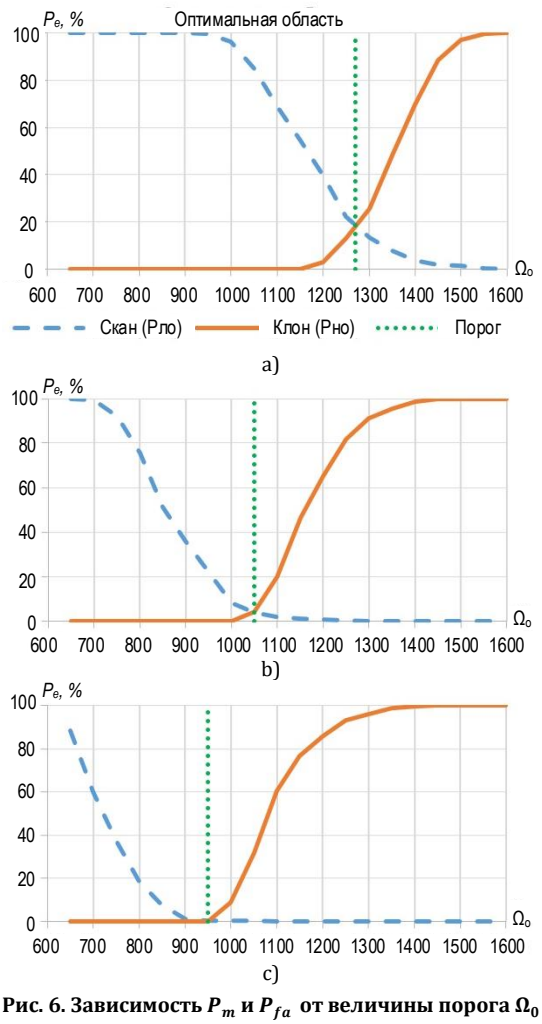


Рис. 6. Зависимость P_m и P_{fa} от величины порога Ω_0 при разрешении сканирования 200 dpi (a), 300 dpi (b) и 600 dpi (c)

Fig. 6. Dependence of P_m u P_{fa} against Ω_0 for Scanning Resolution 200 dpi (a), 300 dpi (b) and 600 dpi (c)

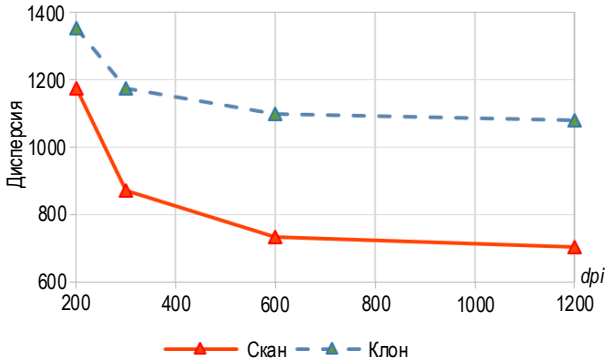


Рис. 7. Зависимость дисперсии шума от разрешения сканирования

Fig. 7. Dependence of the Noise Dispersion on the Scanning Resolution

Для наглядного сравнения экспериментов с проведенными ранее теоретическими исследованиями, проведем расчет полной вероятности ошибки P_e для оптимального порога при различных длинах тестируемой последовательности N (рисунок 8).

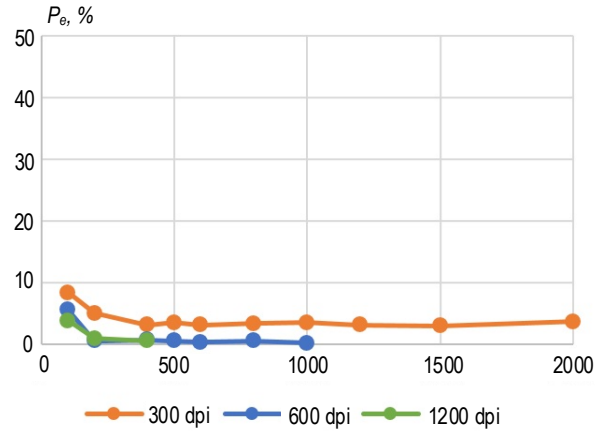


Рис. 8. Экспериментальная зависимость вероятности ошибки от длины тестируемой области *N* при различных значениях dpi

Fig. 8. Experimental Dependence of the Error Probability on the Length of the Tested Area N at Different dpi Values

Для удобного сравнения теоретических и экспериментальных данных отобразим результаты на одном увеличенном графике (рисунок 9).

Привести значение dpi к теоретическому значению r не представляется возможным, однако можно заметить, что графики для разрешения в 600 и 1200 dpi сопоставимы друг с другом и соотносятся с теоретическими для значений $r = 2$ и $r = 4$. А при разрешении в 300 dpi, график согласуется с теоретическим значением $r \approx 6$. Однако он не стремится к 0 %, а выпрямляется при вероятности ошибки в 4 %. Это обусловлено тем, что при малом dpi второстепенные факторы, влияющие на вероятность ошибки, начинают оказывать заметное влияние.

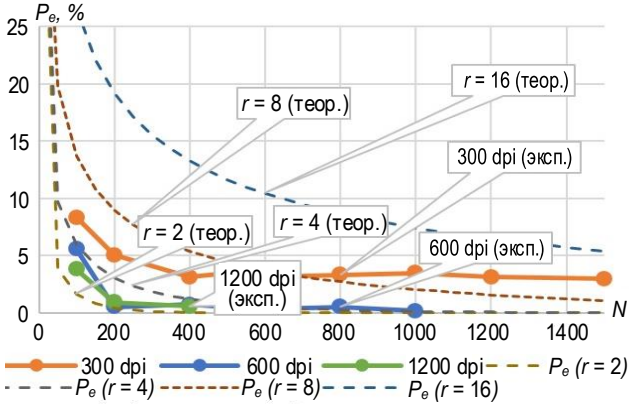


Рис. 9. Сравнение зависимости от *N* теоретической и экспериментальной вероятности ошибки

Fig. 9. Comparison of Theoretical and Experimental Values of Error Probability against the Length N

Заключение

В настоящей работе был экспериментально исследован метод обнаружения атаки клонирования, предложенный ранее в работе [6]. Проведены эксперименты для оценки его эффективности, включающие печать и последующее сканирование изображений, а также аналогичные операции с поддельными сертификатами. Для указанных экспериментов проведена оценка вероятности пропуска, а также ложного обнаружения атаки клонирования. В результате исследования определено подходящее пороговое значение уровня шумов, которое обеспечивает наименьшую вероятность ошибки.

На основе полученных статистических данных можно сделать вывод о том, что предлагаемый метод может быть успешно применен для защиты от атаки клонирования бумажных сертификатов. Стоит заметить, что для эффективной работы метода требуется определение порога уровня шумов, который, в свою очередь, зависит от параметров используемого оборудования, качества имеющихся материалов и параметров вложения, таких как физический размер сертификата.

Для дальнейшего исследования применимости метода в более широком диапазоне практических задач (например, защита товарных этикеток, печатных документов и т. п.) требуется проведение дополнительных экспериментов по определению оптимального порогового значения уровня шумов для различных параметров системы и условий ее использования, таких как качество бумаги, повреждение или загрязнение сертификата на момент сканирования, а также параметры используемого оборудования, и др.

Список источников

1. Коржик В.И., Анфиногенов С.О., Кочкарёв А.И., Федянин И.А., Жувикин А.Г., Флаксман Д.А. Цифровая стеганография и цифровые водяные знаки. Часть 2. Цифровые водяные знаки. СПб: СПбГУТ, 2017. 198 с.

2. Tkachenko I. Generation and analysis of graphical codes using textured patterns for printed document authentication. D.Sc Thesis. Montpellier: Université de Montpellier, 2015.
3. Nguyen H.P., Delahaies A., Retraint F., Nguyen D.H., Pic M., Morain-Nicolier F. A watermarking technique to secure printed QR codes using a statistical test // Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP, Montreal, Canada, 14–16 November 2017). IEEE, 2017. PP. 288–292. DOI:10.1109/GlobalSIP.2017.8308650
4. Chen C., Li M., Ferreira A., Huang J., Cai R. A Copy-Proof Scheme Based on the Spectral and Spatial Barcoding Channel Models // IEEE Transactions on Information Forensics and Security. 2019. Vol. 15. PP. 1056–1071. DOI:10.1109/TIFS.2019.2934861
5. Taran O., Bonev S., Voloshynovskiy S. Clonability of Anti-counterfeiting Printable Graphical Codes: A Machine Learning Approach // Proceedings of the ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP, Brighton, UK, 12–17 May 2019). 2019. PP. 2482–2486. DOI:10.1109/ICASSP.2019.8682967
6. Коржик В.И., Старостин В.С., Флаксман Д.А. Разработка метода использования цифровых водяных знаков для защиты от атаки клонирования бумажных сертификатов // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 79–84. DOI:10.31854/1813-324X-2021-7-2-79-84
7. Korzhik V., Starostin V., Yakovlev V., Flaksman D., Bukshin I., Izotov B. Digital Watermarking System for Hard Cover Objects Against Cloning Attacks // Proceedings of the XXth Conference of Open Innovations Association FRUCT (Oulu, Finland, . 27–29 October 2021). IEEE, 2021. PP. 79–85. DOI:10.23919/FRUCT53335.2021.9599967
8. Solomon C., Breckin T. *Fundamentals of digital signal processing*. Wiley, 2011.
9. Korzhik V., Starostin V., Yakovlev V., Flaksman D. Digital Watermark System with an Ability of its Extraction from Hard Copies of Data // Труды учебных заведений связи. 2019. Т. 5. № 3. С. 75–85. DOI:10.31854/1813-324X-2019-5-3-75

References

1. Korzhik V.I., Anfinogenov S.O., Kochkaryov A.I., Fedyanin I.A., Zhuvikin A.G., Flaksman D.A. *Digital steganography and digital watermarks. Part 2. Digital watermarks*. St. Petersburg: The Bonch-Bruevich Saint Petersburg State University of Telecommunications Publ.; 2017. 198 p.
2. Tkachenko I. *Generation and analysis of graphical codes using textured patterns for printed document authentication*. D.Sc Thesis. Montpellier: Université de Montpellier; 2015.
3. Nguyen H.P., Delahaies A., Retraint F., Nguyen D.H., Pic M., Morain-Nicolier F. A watermarking technique to secure printed QR codes using a statistical test. *Proceedings of the IEEE Global Conference on Signal and Information Processing, GlobalSIP, 14–16 November 2017, Montreal, Canada.* IEEE; 2017. p.288–292. DOI:10.1109/GlobalSIP.2017.8308650
4. Chen C., Li M., Ferreira A., Huang J., Cai R. A Copy-Proof Scheme Based on the Spectral and Spatial Barcoding Channel Models. *IEEE Transactions on Information Forensics and Security*. 2019;15:1056–1071. DOI:10.1109/TIFS.2019.2934861
5. Taran O., Bonev S., Voloshynovskiy S. Clonability of Anti-counterfeiting Printable Graphical Codes: A Machine Learning Approach. *Proceedings of the ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing, UK, 12–17 May 2019).* IEEE; 2019. p.2482–2486. DOI:10.1109/ICASSP.2019.8682967
6. Korzhik V., Starostin V., Flaksman D. Elaboration of Digital Watermarking Method for a Protection of Cloning Attack on Paper Certificates. *Proceedings of Telecommun. Univ*. 2021;7(2):79–84. DOI:10.31854/1813-324X-2021-7-2-79-84
7. Korzhik V., Starostin V., Yakovlev V., Flaksman D., Bukshin I., Izotov B. Digital Watermarking System for Hard Cover Objects Against Cloning Attacks. *Proceedings of the XXth Conference of Open Innovations Association FRUCT*. IEEE; 2021. p 79–85. DOI:10.23919/FRUCT53335.2021.9599967
8. Solomon C., Breckin T. *Fundamentals of digital signal processing*. Wiley, 2011.
9. Korzhik V., Flaksman D. Digital Watermark System with an Ability of its Extraction from Hard Copies of Data. Proceedings of Telecommun. Univ. 2019;5(3):75–85. 2019. DOI:10.31854/1813-324X-2019-5-3-75

Статья поступила в редакцию 16.05.2023; одобрена после рецензирования 01.07.2023; принята к публикации 02.08.2023.
The article was submitted 16.05.2023; approved after reviewing 01.07.2023; accepted for publication 02.08.2023.

Информация об авторе:

ФЛАКСМАН Дмитрий Алексеевич
программист ООО «Научно-производственное предприятие Новые Технологии Телекоммуникаций»
https://orcid.org/0000-0002-0326-4592

СПбГУТ)))

Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М. А. Бонч-Бруевича

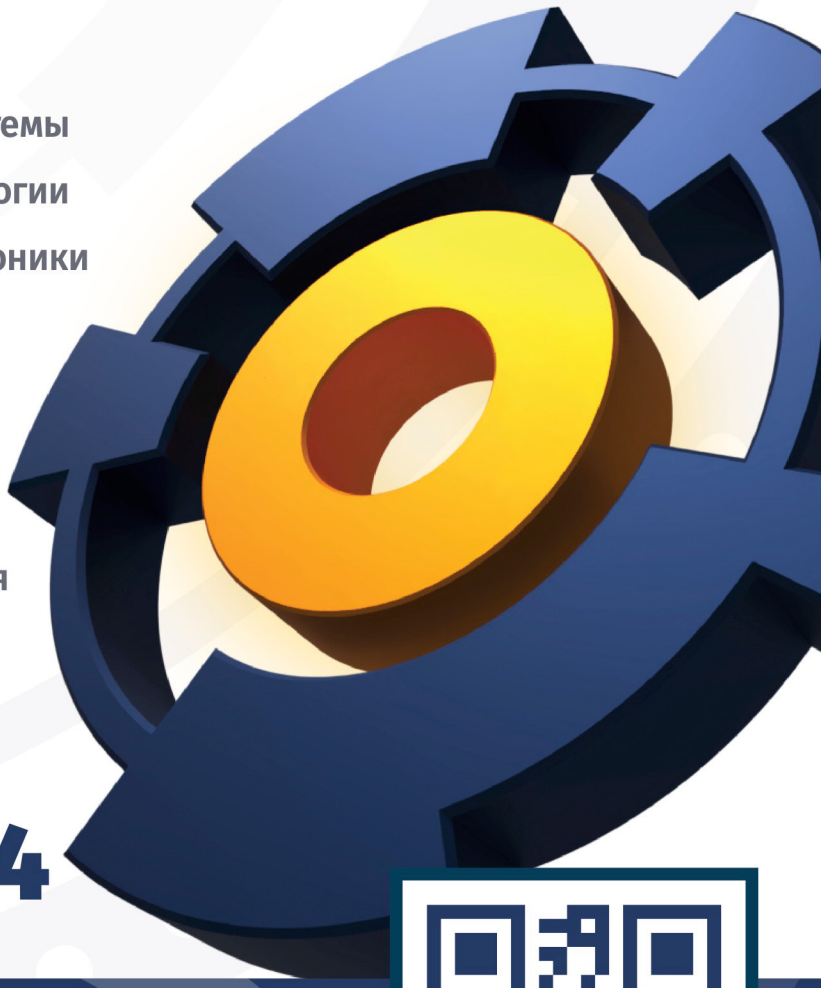
13TH INTERNATIONAL CONFERENCE ON ADVANCED INFOTELECOMMUNICATIONS ICAIT 2024

Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании»

АПИНО ICAIT

ОСНОВНЫЕ НАУЧНЫЕ НАПРАВЛЕНИЯ:

- Радиотехнологии связи
- Инфокоммуникационные сети и системы
- Информационные системы и технологии
- Теоретические основы радиоэлектроники и систем связи
- Цифровая экономика, управление и бизнес-информатика
- Гуманитарные проблемы информационного пространства
- Сети связи специального назначения



27-28 ФЕВРАЛЯ 2024

ПОДРОБНОСТИ НА САЙТЕ КОНФЕРЕНЦИИ

APINO.SPBGUT.RU

22 31 десятилетие науки и технологий

23—26 апреля 2024

СВЯЗЬ

36-я международная выставка «Информационные и коммуникационные технологии»

Экспозиция «Навитех» — «Навигационные системы, технологии и услуги»

www.sviaz-expo.ru

Россия, Москва, ЦВК «ЭКСПОЦЕНТР»

Expo Rating

12+
Реклама

Организатор
ЭКСПОЦЕНТР

При поддержке
МИНПРОМТОРГ РОССИИ

Под патронатом
ТПП РФ

В рамках
Российская неделя высоких технологий
РНВТ

Выходные данные

Товарный знак №929373, правообладатель: Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича»,
191186, Санкт-Петербург, наб. реки Мойки, 61, литера А (RU)
Зарегистрирован в Государственном реестре товарных знаков и знаков обслуживания Российской Федерации 13.03.2023 г. Заявка №2022733914

План издания научной литературы 2023 г., доп. п. 1

Дата выхода в свет	Усл.-печ. л.	Формат	Тираж	Заказ	Свободная цена
27.11.2023	16	$60\times84_{1/8}$	1000 экз.	№ 1540	

Ответственный редактор **Татарникова И.М.**
Выпускающий редактор **Яшугин Д.Н.**
Дизайн: **Коровин В.М.**
Обложка: https://lovepik.com

Адрес СПбГУТ:
19323, Санкт-Петербург, пр. Большевиков, 22/1

Учредитель и издатель:
Федеральное государственное бюджетное образовательное учреждение
высшего образования "Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М.А. Бонч-Бруевича"
E-mail: tuzs@sut.ru Web: tuzs.sut.ru VK: vk.com/spbtuzs

Подписной индекс в Объединенном каталоге "ПРЕССА РОССИИ" - 59983