ISSN 1813-324X (Print)
ISSN 2712-8830 (Online)

ТРУДЫ УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Том 10. № 6
2024

Темы номера:

- Алгоритм синтеза групп кодов в RFID-системе множественного доступа
- Методы пространственной обработки спутниковых навигационных сигналов в частотной области
- Реализация демодулятора сигналов с прямым расширением спектра с использованием методов передискретизации

Vol. 10. Iss. 6
2024

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 10. № 6

Proceedings of Telecommunication Universities

Vol. 10. Iss. 6

Санкт-Петербург

2024

Труды учебных заведений связи. 2024. Т. 10. № 6
ISSN 1813-324X (Print), ISSN 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, статистика
2.3.6. Методы и системы защиты информации, информационная безопасность

Журнал позиционирует себя как научный, в связи с этим его целями являются ознакомление научной общественности (научного сообщества) с результатами оригинальных исследований, выполненных ведущими учеными и специалистами и их коллективами, а также апробация научных результатов, полученных при подготовке кандидатских и докторских диссертаций для повышения качества (уровня) проводимых исследований. Издание ставит перед собой задачу расширения инфокоммуникативного пространства взаимодействия российских и зарубежных ученых. Целевой аудиторией журнала являются ученые и специалисты-практики в области связи и телекоммуникаций и смежных направлениях науки и техники, а также профессорско-преподавательский состав и студенты, обучающиеся по программам аспирантуры, магистратуры, специалитета и бакалавриата профильных вузов и кафедр.

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

Proceedings of Telecommunication Universities. 2024. Vol. 10. Iss. 6
Trudy učebnyh zavedenij svâzi. ISSN 1813-324X (Print), ISSN 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, statistics
2.3.6. Methods and systems of information security, cybersecurity

The journal positions itself as a scientific one, in this regard, its goals are to familiarize the scientific community (scientific community) with the results of original research carried out by leading scientists and specialists and their teams, as well as approbation of scientific results obtained in the preparation of candidate and doctoral dissertations to improve the quality (level) of ongoing research. The publication sets itself the task of expanding the infocommunicative space of interaction between Russian and foreign scientists. The target audience of the journal are scientists and practitioners in the field of communications & telecommunications and related fields of science & technology, as well as faculty and students enrolled in postgraduate, master's, specialisation and bachelor's programs of profiled universities and departments.

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

Editorial Board

РЕГИСТРАЦИОННАЯ ИНФОРМАЦИЯ / REGISTRATION INFORMATION

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.)

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

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)

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

Товарный знак № 929373.

Правообладатель:

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

191186, Санкт-Петербург, наб. реки Мойки, 61, литера А

Trademark No. 929373.

Copyright holder:

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

191186, St. Petersburg, emb. Moika River, 61, letter A

КОНТАКТНАЯ ИНФОРМАЦИЯ / CONTACT INFORMATION

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

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

Ответственный редактор **Татарникова И.М.**
Выпускающий редактор **Яшугин Д.Н.**
Дизайн: **Коровин В.М.**

Executive Editor **Tatarnikova I.M.**
Commissioning Editor **Yashugin D.N.**
Design: **Korovin V.M.**

ВЫХОДНЫЕ ДАННЫЕ / IMPRINT

Дата выхода в свет: 28.12.2024
Тираж: 1000 экз. Цена свободная.

Отпечатано в типографии Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича»

Release date: 28.12.2024
Circulation: 1000 copies. Free price.

Printed in the printing office Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruevich Saint-Petersburg State University of Telecommunications»

СОДЕРЖАНИЕ

CONTENTS

DOI:10.31854/1813-324X-2024-10-6

Научная статья
УДК 621.372.632
https://doi.org/10.31854/1813-324X-2024-10-6-7-18

Реализация демодулятора сигналов с прямым расширением спектра с использованием методов передискретизации

Ефим Александрович Брусин, brusin.ea@sut.ru

Институт радионавигации и времени АО «Обуховский завод»,
Санкт-Петербург, 192012, Российская Федерация
Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. *В последние годы широкое распространение в системах связи и навигации находят сигналы с прямым расширением спектра. В частности, эти сигналы превалируют в современных системах спутниковой навигации и используются в системах связи с кодовым разделением каналов. Поэтому задачи построения демодуляторов сигналов с прямым расширением спектра приобретают ключевое значение. Особую значимость при построении демодуляторов приобретает проблема их демодуляторов по скорости следования чипов.*

Цель исследования *состоит в том, чтобы предложить структуру демодулятора, ориентированную на решение указанной проблемы. Исследование основано на* ***методах*** *компьютерного моделирования.*

Решение. *В работе переложен подход к построению демодуляторов сигналов с прямым расширением спектра, основанный на современных методах цифровой обработки сигналов. Показано, что главным преимуществом предлагаемого подхода является возможность перестройки демодулятора по чиповой скорости. На основании полученных результатов предложена схема демодулятора сигналов с прямым расширением спектра, использующего методы передискретизации. Передискретизация сигнала, в свою очередь, реализуется на основе полиномиальной интерполяции с использованием полиномов Лагранжа. Предложена структура передискретизатора, подобная структуре интерполирующего фильтра с конечной импульсной характеристикой. Представленные результаты моделирования показывают эффективность предложенного подхода.*

Новизна. *Представляется, что распространенные в настоящее время подходы к реализации демодуляторов сигналов с прямым расширением спектра в части синхронизации по задержке не отвечают в достаточной степени современным требованиям. Построение схемы синхронизации по задержке на основе передискретизации практически не обсуждается в известных работах. В тоже время современные методы и устройства цифровой обработки сигналов позволяют обеспечить эффективную аппаратную реализацию рассматриваемой схемы. В этой связи предложенный в работе подход к построению демодуляторов представляется весьма актуальным.*

Значимость. *Результаты работы могут использоваться при построении демодуляторов сигналов с прямым расширением спектра для широкого круга систем связи и навигации. Структура с асинхронной дискретизацией, предложенная в работе, весьма перспективна особенно для демодуляторов, перестраиваемых по чиповой скорости.*

Ключевые слова: *прямое расширение спектра, демодулятор, асинхронная дискретизация, полиномиальная интерполяция, передискретизатор*

Ссылка для цитирования: Брусин Е.А. Реализация демодулятора сигналов с прямым расширением спектра с использованием методов передискретизации // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 7–18. DOI:10.31854/1813-324X-2024-10-6-7-18. EDN:JVRUQG

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-7-18

The Resampling Methods Direct Sequence Spread Spectrum Signal's Demodulator Implementation

Efim A. Brusin, brusin.ea@sut.ru

Institute of Radio Navigation and Time JSC «Obukhov Plant»,
St. Petersburg, 192012, Russian Federation
The Bonch-Bruevich Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance. *The direct spread spectrum signals are widely used in navigation and communication systems recently. These signals prevail in modern satellite navigation systems and are used in various communication systems with code division multiplexing in particularly. In this regard, the tasks of building direct spread spectrum signals' demodulators have the key importance. Mach importance in the construction of demodulators is the problem chip rate variability.*
The purpose of the study *is to propose a demodulator structure focused on solving this problem.*
Methods. *The research is based on computer modeling methods.*
Decision. *The paper proposes an approach to the construction of the direct spread spectrum signal's demodulators based on modern methods of digital signal processing. It is shown that the main advantage of the proposed approach is the possibility of rebuilding the variable chip rate demodulators. Based on the results obtained, a scheme for the direct spread spectrum signals demodulator using resampling methods is proposed. Resampling, in turn, is implemented on the basis of polynomial interpolation using Lagrange polynomials. The structure of the resampler is proposed, similar to the structure of an interpolating filter with a finite impulse response. The presented simulation results show the effectiveness of the proposed approach.*
Novelty*. It seems that the currently common methods of implementing direct spread spectrum signal in terms of delay synchronization do not sufficiently meet modern requirements. The implementation of delay synchronization schemes based on resampling is practically not discussed in well-known works. At the same time, modern methods and devices of digital signal processing make it possible to ensure an effective hardware implementation of the scheme in question. In this context, the approach proposed in the paper to the construction of demodulators seems to be very relevant.*
Significance. *The results of the work can be used in the construction with direct spread spectrum signals' demodulators for a wide range of communication and navigation systems. The synchronous sampling structure proposed in this paper is very promising, especially for variable chip rate demodulators.*

Key words: *direct sequence spread spectrum, demodulator, asynchronous sampling, polynomial interpolation, resampler*

For citation: Brusin E.A. The Resampling Methods Direct Sequence Spread Spectrum Signal's Demodulator Implementation. *Proceedings of Telecommunication Universities.* 2024;10(6):7–18. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-7-18. EDN:JVRUQG

1. Введение

Одной из ключевых проблем, возникающих при реализации цифрового демодулятора, является проблема синхронизации по задержке или, иначе говоря, синхронизация по границам символов (символьной синхронизации). Естественно, задача синхронизации в цифровых демодуляторах решается с учетом дискретизации принимаемого сигнала. Вопросы дискретизации, в частности, обсуждались в работах [1] и [2] в свете реализации методов передачи и приема сигналов в цифровых модемах.

В части реализации схемы дискретизации сигнала цифрового демодулятора можно выделить

два подхода: синхронная и асинхронная дискретизация. Структура демодулятора с синхронной дискретизацией представлена на рисунке 1a. Принимаемый сигнал поступает на вход схемы аналоговой обработки, а с выхода последней – на вход аналого-цифрового преобразователя (АЦП). Схема цифровой обработки формирует сигнал подстройки частоты управляемого генератора, формирующего сигнал частоты дискретизации. То есть синхронная дискретизация предполагает непосредственное управление частотой дискретизации АЦП.

Структура демодулятора с асинхронной дискретизацией представлена на рисунке 1b. В данном случае частота дискретизации АЦП формируется независимым генератором. А уже схема цифровой обработки решает вопросы синхронизации. Основным достоинством схемы с синхронной дискретизацией является простота реализации в части цифровой обработки сигнала. Проблемы возникают при решении задачи перестройки по скорости передачи данных.

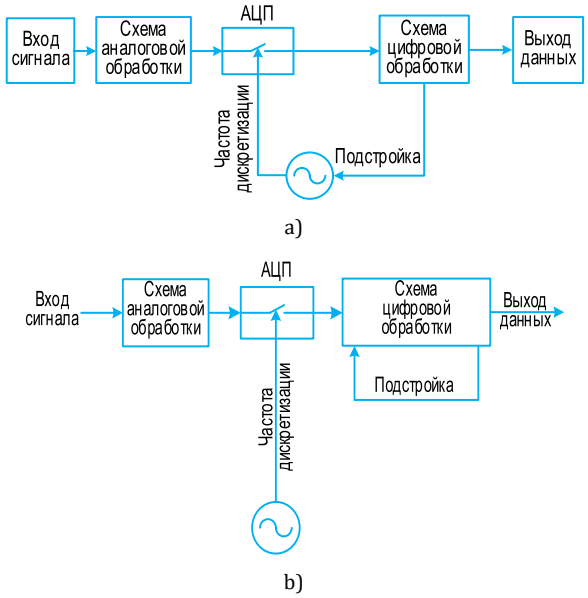


Рис. 1. Демодулятор с синхронной (a) и асинхронной (b) дискретизацией

Fig. 1. Synchronized (a) and No Synchronized (b) Sampling Demodulator

Для схемы с синхронной дискретизацией: перестройка по скорости передачи данных зачастую требует изменения частоты дискретизации и, как правило, – подстройку полосы фильтров, в функции которых входит ограничение полосы сигнала на входе АЦП для предотвращения эффекта наложения спектров. Схема с асинхронной дискретизацией свободна от этого недостатка, так как частота дискретизации может быть выбрана таким образом, чтобы избежать эффекта наложения спектров во всем диапазоне скоростей передачи данных.

При построении демодуляторов сигналов с прямым расширением спектра, в частности для аппаратуры потребителя глобальных навигационных спутниковых систем, как правило, используются методы синхронной дискретизации [3, 4]. Структура демодулятора сигналов с прямым расширением спектра, использующего подход с асинхронной дискретизацией, представлена на портале ComBlocks https://www.comblock.com/com1518soft.html. Однако демодулятор здесь описан достаточно поверхностно, не обсуждается проблема перестройки по чиповой скорости. В то же время для систем связи, навигации и систем сличения шкал времени, использующих сигналы с прямым расширением спектра, частота следования чипов может изменяться в значительном диапазоне[1, 2] [5]. То есть подходы к построению демодулятора сигналов с прямым расширением спектра в целом достаточно известны, но реализация демодулятора с использованием схем асинхронной дискретизации практически не обсуждается. В то же время с развитием технологий цифровой обработки сигналов методы асинхронной дискретизации широко используются. В частности, такие подходы представлены в работах [1, 2, 6–8]. На рисунке 2 демонстрируется упрощенная схема символьной синхронизации демодулятора с асинхронной дискретизацией.

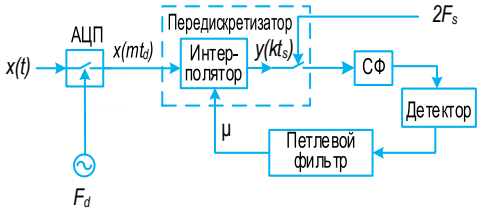


Рис. 2. Асинхронная дискретизация. Схема символьной синхронизации

Fig. 2. No Synchronized Sampling. The Symbol Synchronization Scheme

Основой схемы является интерполятор. На вход интерполятора поступают отсчеты сигнала с выхода АЦП. На выходе интерполятора формируются отсчеты сигнала на частоте, равной удвоенной частоте следования символов сигнала. F_d – частота дискретизации (см. рисунок 2). Для систем без расширения спектра F_S – частота следования символов, а с прямым расширением спектра F_S – частота следования чипов. Сигнал с выхода согласованного фильтр (СФ) с частотой следования отсчетов $2F_S$ поступает на вход схемы детектора петли синхронизации по задержке, а с выхода последнего – на

[1]Интерфейсный контрольный документ. Радиосигналы и состав цифровой информации функционального дополнения системы ГЛОНАСС системы дифференциальной коррекции и мониторинга (редакция 1). URL: https://sdcm.ru (дата обращения 22.11.2023)

[2] Российские космические системы. ГЛОНАСС. Интерфейсный контрольный документ. URL: https://russianspacesystems.ru/bussines/navigation/glonass/interfeysnyy-kontrolnyy-document (дата обращения 09.09.2024)

вход петлевого фильтра. На выходе петлевого фильтра формируются отсчеты сигнала управления фазой интерполятора – μ. Собственно, интерполятор в совокупности со схемой преобразования частоты дискретизации образуют передискретизатор.

Коэффициент передискретизации k_p определяется следующим образом:

$$k_p = \frac{2F_S}{F_d}.$$

Согласованный фильтр в схеме для сигнала с прямым расширением спектра – фильтр, реализованный на основе коррелятора с заданной расширяющей последовательностью. Собственно, структура демодулятора сигнала с прямым расширением спектра известна. Однако представляется, что использование традиционных подходов к решению задачи синхронизации по задержке на основе синхронной дискретизации затрудняет перестройку демодулятора по чиповой скорости. В этом смысле интерес представляет реализация демодулятора сигнала с прямым расширением спектра с использованием методов передискретизации. Таким образом, основными задачами представляемой статьи являются построение демодулятора с использованием передискретизации и исследование предложенного демодулятора методами компьютерного моделирования.

2. Реализация демодулятора

Передискретизация сигнала, как правило, реализуется на основе полиномиальной интерполяции. Полиномиальный интерполятор был предложен в работе Фарроу [9], где рассматривалась реализация элемента с дробной задержкой на основе интерполяции с использованием полиномов Лагранжа. При реализации передискретизатора возможен ряд подходов. Структура Фарроу реализуется путем прямого вычисления интерполированных отсчетов. Интерполятор может быть также реализован на основе структуры, подобной фильтру с конечной импульсной характеристикой (КИХ-фильтру). Рассматриваемый подход иллюстрирует схема формирования передаваемого сигнала, представленная на рисунке 3. На вход схемы поступают отсчеты сигнала x_i на удвоенной чиповой частоте. Для двухпозиционной фазовой модуляции (ФМ-2) логическая единица передается как +1, ноль как –1. Схема включает в себя накапливающий сумматор, формирующий отсчеты сигнала фазы передискретизации μ_k. Отсчеты μ_{k-1} задержанной фазы складываются с коэффициентом передискретизации.

Если $\mu_k > 1$, то отсчеты сигнала x_i записываются в буфер передискретизатора. При этом из текущего значения фазы передискретизатора вычитается единица: $\mu_k = \mu_k - 1$.

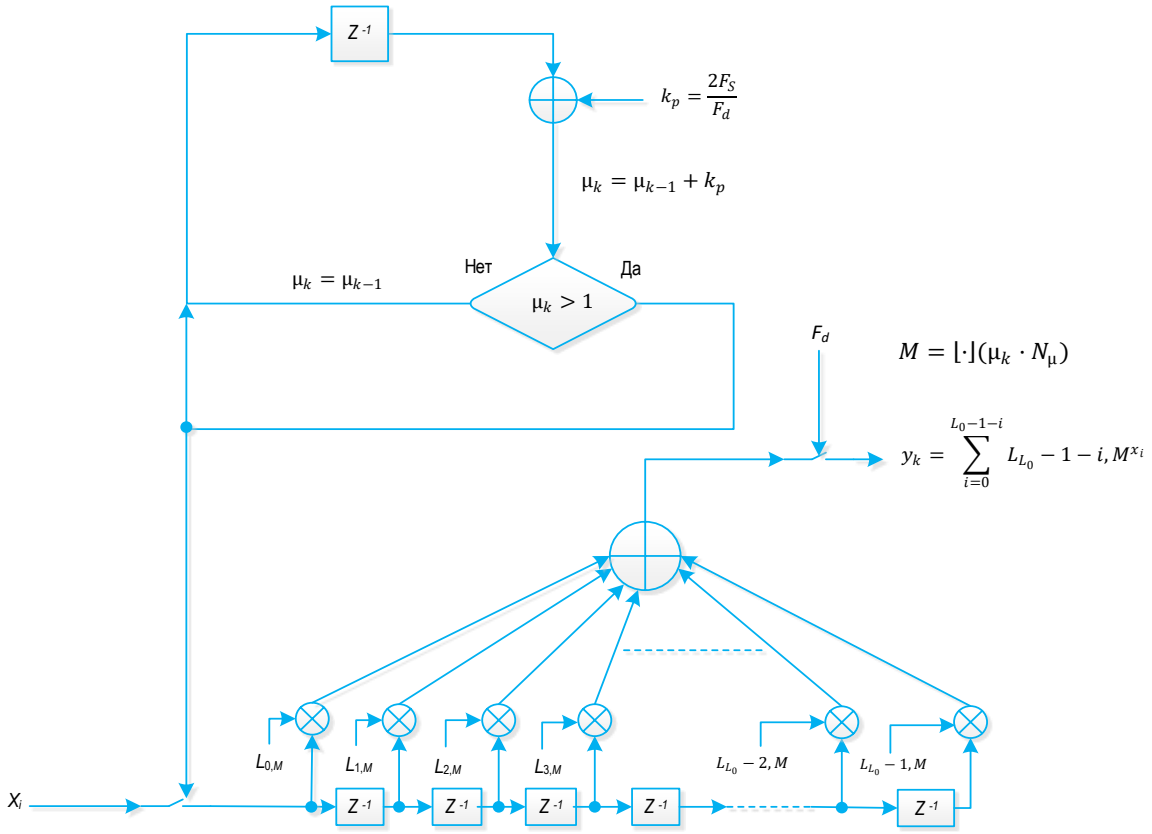


Рис. 3. Передискретизатор на передачу

Fig. 3. The Transmitter Resampler

Отсчеты на выходе передискретизатора y_k вычисляются как свертка хранящихся в таблице отсчетов полиномов с отсчетами сигнала:

$$y_k = \sum_{i=0}^{L_0-1-i} L_{L_0-1-i,M}\, x_i,$$

где $L_{L_0-1-i,M}$ – отсчеты интерполирующих полиномов Лагранжа; L_0 – размерность полиномов; M – фаза передискретизатора; N_μ – величина, связанная с размерностью таблицы; $\lfloor\cdot\rfloor$ – означает усечение до целого.

В качестве иллюстрации работы схемы на рисунках 4 и 5 показаны отсчеты сигналов передискретизатора. В рассматриваемом случае F_S = 20 МГц, F_d = 140 МГц. Коэффициент передискретизации k_P равен 2/7. L_0 = 8, N_μ = 1023. Соответственно, число фаз передискретизатора равно 1024, размерность таблицы полиномов – 8 на 1024.

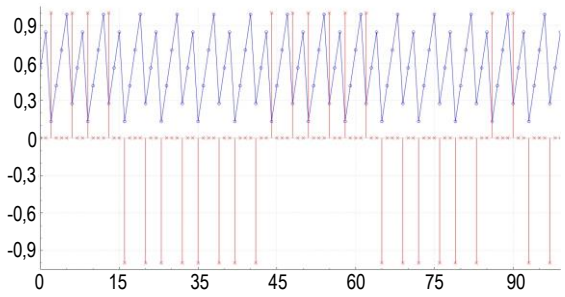


Рис. 4. Отсчеты сигналов μ_k и x_i

Fig. 4. μ_k and x_i Signals Samples

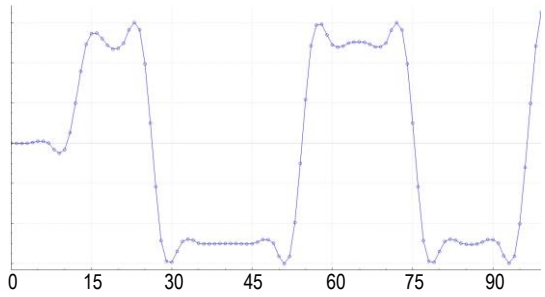


Рис. 5. Отсчеты сигнала на выходе передискретизатора

Fig. 5. Resampler Output Signals Samples

На рисунке 4 представлены отсчеты сигналов фазы передискретизатора и отсчеты сигнала x_i. Сигнал фазы передискретизатора представляет собой «пилу», наклон которой определяется значением коэффициента передискретизации. Заметим, что отсчеты входного сигнала записываются в буфер передискретизатора при переходе фазы через единицу. Вычисление отсчетов y_k производится синхронно с выходной частотой дискретизации F_d в соответствии с текущим значением фазы:

$$M = \lfloor\cdot\rfloor(1023 \cdot \mu_k).$$

Соответствующие выходные отсчеты передискретизатора представлены на рисунке 5.

Предлагаемая структура демодулятора сигнала с прямым расширением спектра, построенного на основе методов передискретизации, представлена на рисунке 6. Принимаемый сигнал поступает на вход схемы комплексного переноса. С выхода последнего – на вход фильтров нижних частот (ФНЧ), в функции которых входит удаление побочных продуктов переноса. Сигналы с выходов указанных ФНЧ поступают на входы каскадов фильтров, предназначенных для уменьшения частоты дискретизации – децимации сигнала. Как правило, децимация осуществляется с использованием CIC-фильтров (CIC, *аббр. от англ.* Cascaded Integral-Comb Filters) [10]. CIC-фильтры обеспечивают децимацию сигнала с коэффициентом, равным степени двойки (2^K).

Таким образом, частота дискретизации на выходе CIC-фильтров определяется по выражению:

$$F_d = F_0/2^K,$$

где F_0 – частота дискретизации принимаемого сигнала; 2^K – коэффициент децимации.

Сигналы квадратур с частотой дискретизации F_d поступают на входы следующих ФНЧ. В функции указанных ФНЧ входит предотвращение эффекта наложения спектров при реализации процедуры передискретизации на приеме. Сигналы S_I и S_Q с выходов ФНЧ поступают на входы схем передискретизаторов квадратурных каналов. Структура передискретизатора на примере схемы передискретизации синфазного канала представлена на рисунке 7. Отсчеты сигнала S_I записываются в буфер передискретизатора с частотой дискретизации F_d. Основой передискретизатора является целочисленный накапливающий сумматор, формирующий пилообразный сигнал.

Старшие 10 разрядов накапливающего сумматора формируют адрес для обращения в таблицу полиномов Лагранжа (ROM, *аббр. от англ.* Read Only Memory, MSB, *аббр. от англ*. Most Significant Bit на рисунке 7).

Как только накапливающий сумматор передискретизатора переходит через максимальное значение (wrap around), вычисляются выходные отсчеты передискретизатора. Момент wrap around определяет изменение из единицы в ноль старшего разряда накапливающего сумматора. Для иллюстрации работы передискретизатора на рисунке 8 представлены отсчеты старшего разряда накапливающего сумматора и отсчеты собственно сумматора. Старший разряд принимает значения 1 и –1. Отсчеты накапливающего сумматора изменяются от 0 до 1 в соответствии со значением коэффициента μ_k. В рассматриваемом случае F_S = 20 МГц, F_d = 140 МГц. В итоге передискретизатор вычисляет свертку отсчетов сигнала с соответствующим набором коэффициентов интерполирующего полинома Лагранжа: $X = \sum_{n=0}^{7} L[k][M]S_I(k)$.

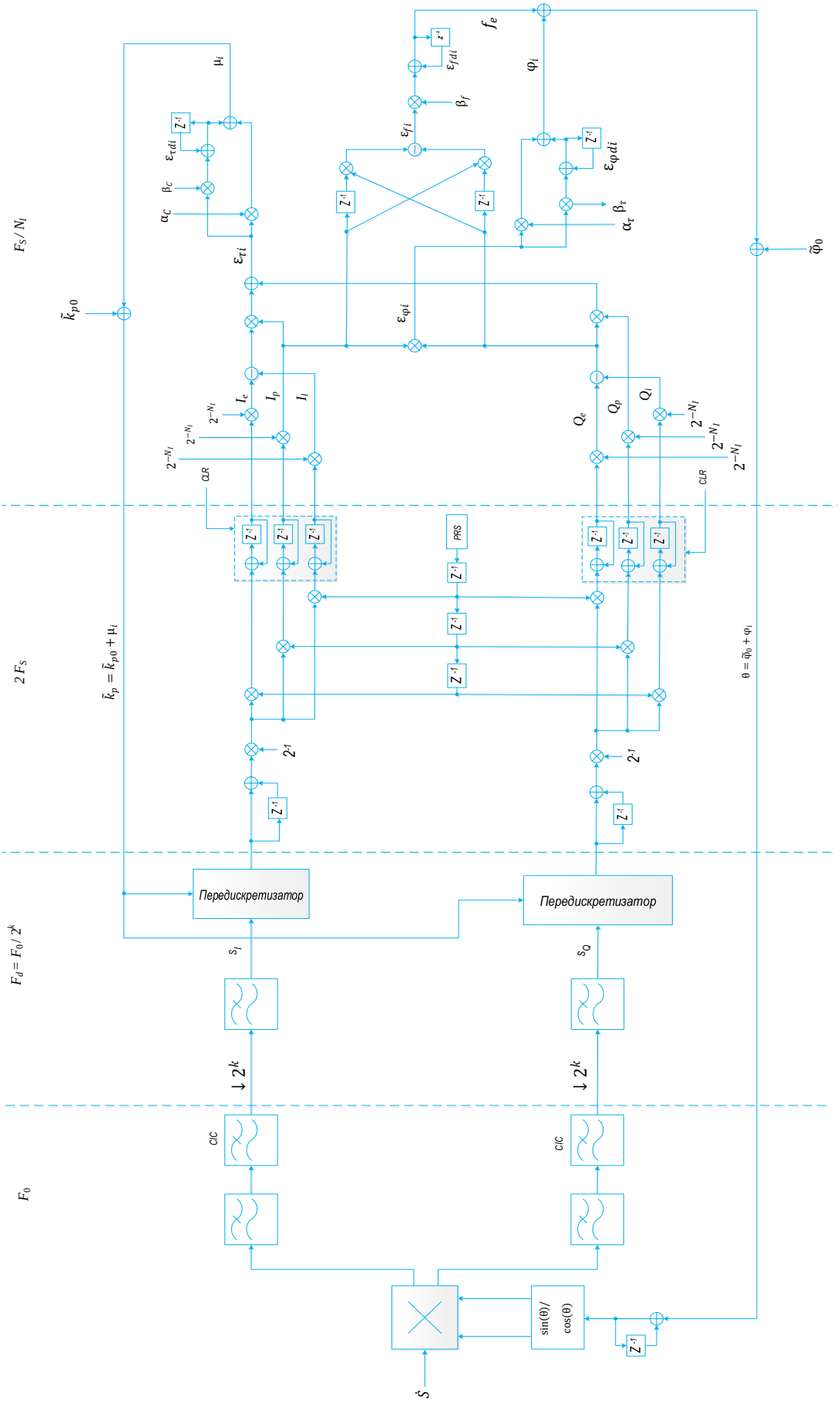


Рис. 6. Демодулятор сигнала
Fig. 6. The Signal Demodulator

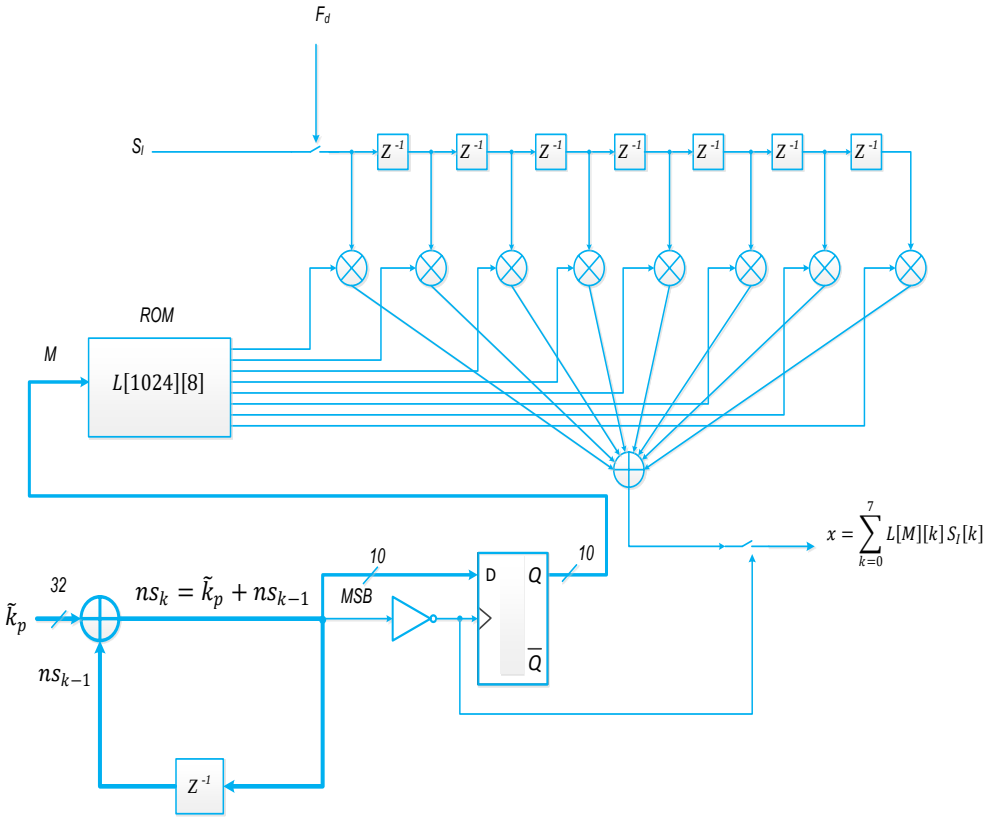


Рис. 7. Передискретизатор синфазного канала

Fig. 7. The In-Phase Channel Resampler

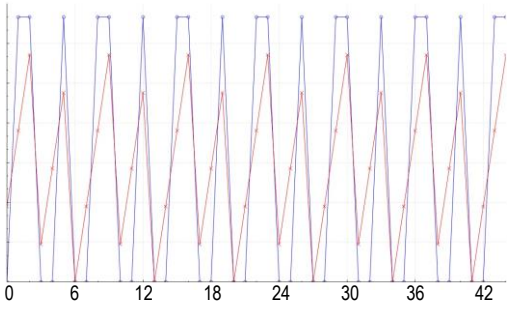


Рис. 8. Отсчеты накапливающего сумматора – *x*, старший разряд накапливающего сумматора – *o*

Fig. 8. Accumulator Samples – x, Accumulator Most Significant Bit – o

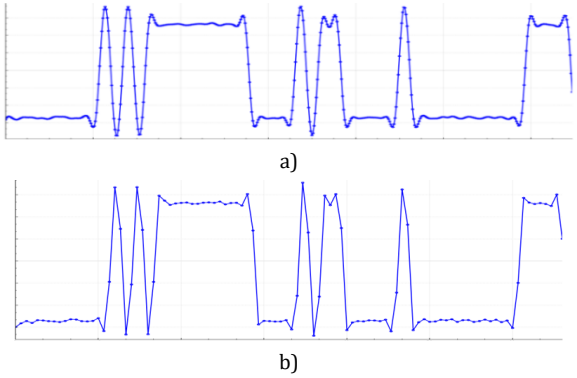


Рис. 9. Отсчеты сигнала на входе (a) и выходе (b) передискретизатора

Fig. 9. Resampler Input (a) and Output (b) Signals Samples

На рисунке 9 представлены отсчеты сигналов на входе и выходе передискретизатора.

В демодуляторе присутствуют петли фазовой и частотной автоподстройки [11]. В качестве детектора сигнала ошибки в петли подстройки по несущей частоте используется фазовый детектор с перемножением сигналов I_p и Q_p:

$$\varepsilon_\varphi = I_p Q_p.$$

Сигнал ошибки частотного детектора вычисляется следующим образом:

$$\varepsilon_f = I_p(n)Q_p(n-1) - Q_p(n)I_p(n-1),$$

где $I_p(n), Q_p(n), Q_p(n-1), I_p(n-1)$ – текущие и задержанные сигналы «prompt».

Получаемые сигналы ошибок поступают на входы петлевых фильтров.

Для реализации петли по задержке демодулятор включает в себя согласованные фильтра квадратурных каналов: «early» – с опережением на пол чипа, «late» – с запаздыванием на пол чипа (PRS на рисунке 6 – генератор заданной последовательности). Детектор ошибки петли по задержке реализован по схеме типа «early–late» [3, 4].

Сигнал ошибки, определяемый как:

$$\varepsilon_\tau = I_p(I_l - I_e) + Q_p(Q_l - Q_e),$$

поступает на вход петлевого фильтра.

Сигнал μ_i с выхода петлевого фильтра управляет коэффициентом передискретизации:

$$\tilde{k}_p = \tilde{k}_{p0} + \mu,$$

где $\tilde{k}_{p0} = \lfloor \cdot \rfloor \left((2F_s/F_d)2^{32} \right)$ – номинальный коэффициент передискретизации демодулятора.

Собственно, в реализации подстройки в петле по задержке и состоит основное отличие предложенного подхода. В традиционном демодуляторе петля по задержке управляет генератором, который фактически формирует опорные «early», «late» и «prompt» последовательности. В рассматриваемом демодуляторе петля по задержке управляет коэффициентом передискретизации.

При построении демодулятора с переменной чиповой скоростью фактически предлагается двухступенчатое преобразование частоты дискретизации. Первая ступень соответствует децимации на CIC-фильтре. Коэффициент децимации CIC-фильтров выбирается таким образом, чтобы выполнялось условие: $F_0/2^K \geq 2F_S$.

Передискретизация в демодуляторе обеспечивает преобразование частоты дискретизации с коэффициентом:

$$k_p = 2F_S/(F_0/2^K).$$

В качестве иллюстрации использования предлагаемого подхода рассмотрим реализацию целочисленной модели демодулятора сигналов на примере приема сигнала ФМ-2. Частота следования чипов F_S соответствует 0,5; 1; 2,5; 5; 10; 20 МГц. При формировании расширения использовались укороченные последовательности Голда длиной N_I от 2000 до 80000. Скорость передачи данных $f_b = 250$ Гц. Частота дискретизации принимаемого сигнала $F_0 =$ 280 МГц. Параметры демодулятора сведены в таблицу 1.

ТАБЛИЦА 1. Параметры демодулятора

TABLE 1. Demodulator Parameters

F_S, МГц	Коэффициент Передискретизации k_P	Коэффициент Децимации (2^K)
20	2/7	2
10	2/7	4
5	2/7	8
2,5	2/7	16
1	8/35	32
0,5	8/35	64

Реализация модели демодулятора для $F_S = 20$ МГц иллюстрируют отсчеты сигналов, представленные на рисунках 10–12. Проблемы синхронизации по несущей частоте подробно описаны в работе [11]. Основное внимание уделим синхронизации по задержке. Вхождение демодулятора в синхронизм иллюстрирует рисунок 10, который демонстрирует подстройку коэффициента передискретизации $k_P = \tilde{k}_p/2^{32}$. Заметим, что коэффициент передискретизации «стягивается» к номинальному значению, равному 2/7 (красная линия на рисунке). На рисунке 11 показаны отсчеты выходных сигналов квадратурных каналов I_p и Q_p, а на рисунке 12 – фазовая диаграмма демодулятора в установившемся режиме.

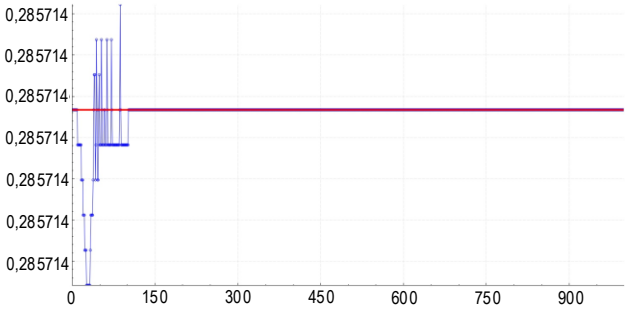


Рис. 10. Коэффициент передискретизации

Fig. 10. Resampling Coefficient

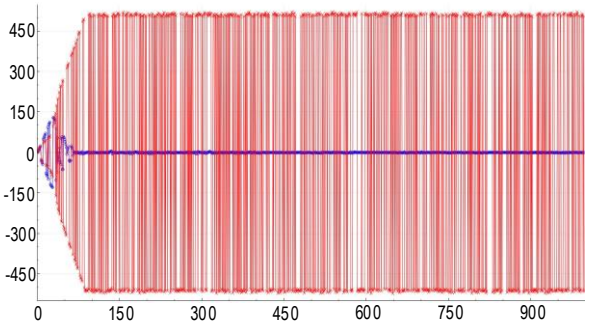


Рис. 11. Отсчеты сигналов: I_p – *x* и Q_p – *o*

Fig. 11. Signal Samples: I_p – x and Q_p – o

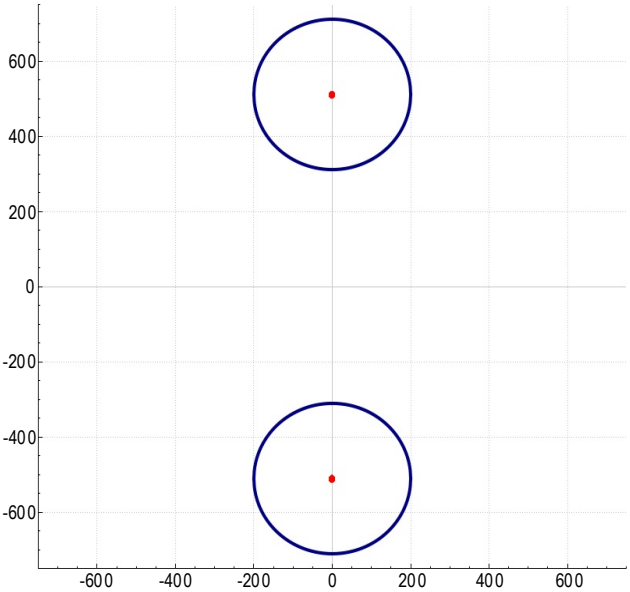


Рис. 12. Фазовая диаграмма принимаемого сигнала

Fig. 12. Receive Signal Phase Diagram

На точки фазовой диаграммы накладываются окружности, которые характеризуют вхождение

демодулятора в синхронизм (захват демодулятора). Окружности задают области захвата демодулятора. Алгоритм определения вхождения демодулятора в захват на основе анализа фазовых диаграмм принимаемого сигнала предложен в [11]. По виду точек на фазовой диаграмме, представленной на рисунке 12, можно судить об энергетических потерях приема. Полученное «сжатие» точек косвенно свидетельствует о малых потерях, обеспечиваемых предложенным демодулятором. На рисунках 13a, 13c, 13e представлены результаты моделирования, иллюстрирующие процесс синхронизации демодулятора для F_S = 20 МГц при E_S/N_0 = –35 дБ.

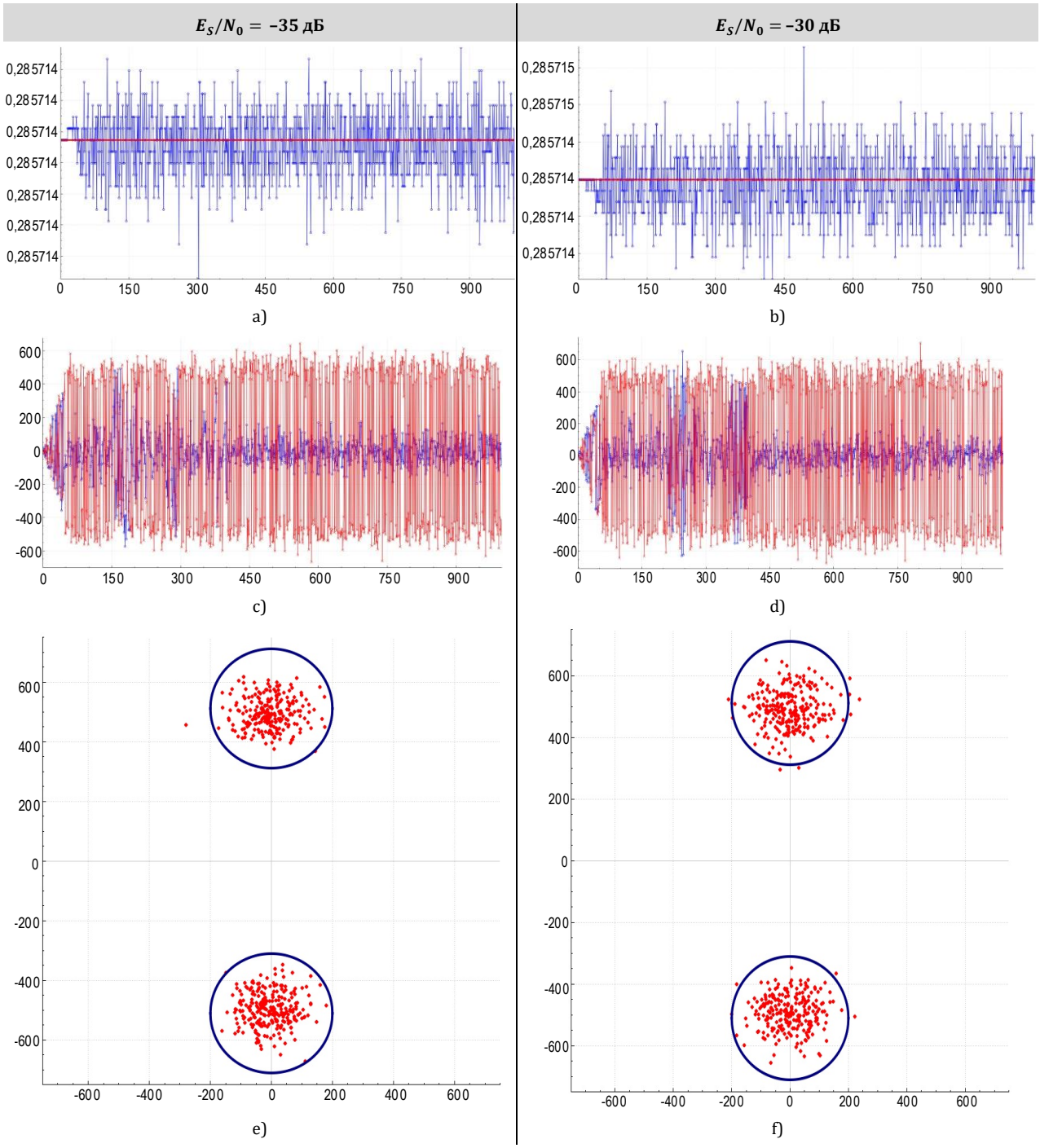


Рис. 13. Результаты моделирования, иллюстрирующие процесс синхронизации демодулятора для F_S = 20 МГц (слева) и F_S = 5 МГц (справа): коэффициенты передискретизации (a, b); отсчеты сигналов I_p – *x* и Q_p – *o* (c, d); фазовые диаграммы принимаемых сигналов (e, f)

Fig. 13. Modeling Results Demodulator Acquisition Process Illustrated for F_S = 20 MHz (left) and F_S = 5 MHz (right): Resampling Coefficient (a, b); Signal Samples I_p – x and Q_p – o (c, d); Receiver Signal Phase Diagrams (e, f)

Рисунок 13a иллюстрирует синхронизацию по задержке. На рисунке 13c представлены отсчеты выходных сигналов демодулятора, а на рисунке 13e – соответствующая фазовая диаграмма принимаемого сигнала.

Для иллюстрации перестройки демодулятора по чиповой скорости на рисунках 13b, 13d, 13f и 14 показаны результаты моделирования для ряда чиповых скоростей: $F_S = 5$, 1 и 0,5 МГц. При $F_S = 5$ МГц, так же, как и для $F_S = 20$ МГц, коэффициент передискретизации стремится к 2/7. При $F_S = 1$ и 0,5 МГц петля подстройки по задержке «стягивает» коэффициент передискретизации к 8/35. Во всех случаях подавляющее большинство точек на фазовых диаграммах лежит внутри областей захвата.

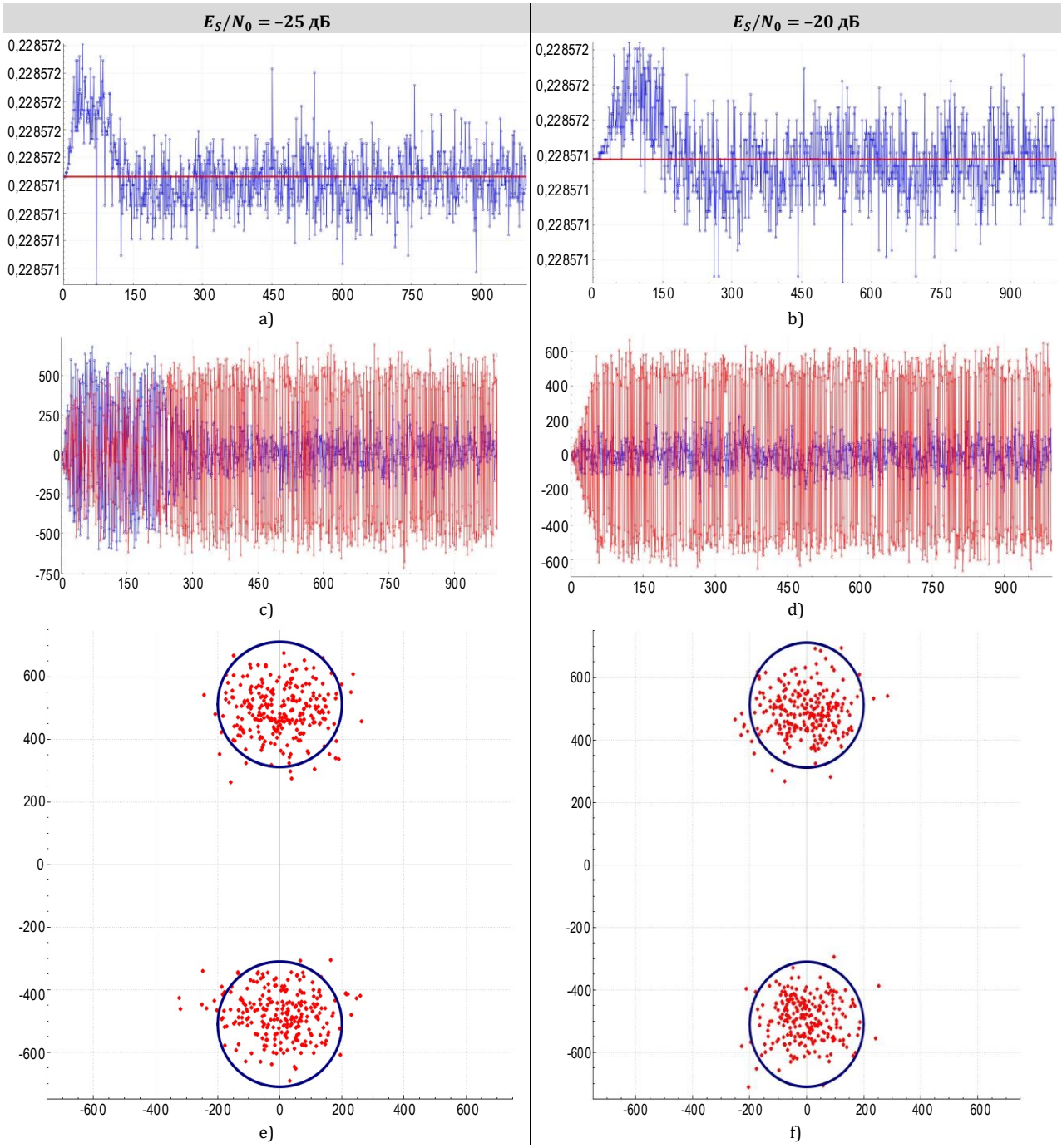


Рис. 14. Результаты моделирования, иллюстрирующие процесс синхронизации демодулятора для $F_S = 1$ МГц (слева) и $F_S = 0{,}5$ МГц (справа): коэффициенты передискретизации (a, b); отсчеты сигналов I_p – *x* и Q_p – *o* (c, d); фазовые диаграммы принимаемых сигналов (e, f)

Fig. 14. Modeling Results Demodulator Acquisition Process Illustrated for $F_S = 1$ MHz (left) and $F_S = 0{,}5$ MHz (right): Resampling Coefficient (a, b); Signal Samples I_p – x and Q_p – o (c, d); Receiver Signal Phase Diagrams (e, f)

3. Выводы

В работе обсуждаются проблемы синхронизации по задержке в демодуляторах сигналов с прямым расширением спектра. При классическом подходе к реализации демодуляторов, основанном на синхронной дискретизации, особые трудности возникают при перестройке по скорости следования чипов.

Предложен подход к построению демодуляторов сигнала с прямым расширением спектра. В работе предложено использовать алгоритмы преобразования частоты дискретизации (передискретизации) с использованием полиномиальной интерполяции.

В настоящее время известны различные варианты построения интерполяторов. Например, на основе схемы Фарроу или же на основе структуры, подобной интерполирующему КИХ-фильтру. В работе предлагается использовать последний подход. Отсчеты интерполирующих полиномов хранятся в табличном виде, а фаза интерполятора выбирает соответствующие отсчеты полиномов. Отсчеты на выходе передискретизатора вычисляются как свертка указанного набора и отсчетов принимаемого сигнала.

Показано, что основное преимущество предложенного подхода проявляется при решении задачи реализации демодулятора с переменной чиповой скоростью. В качестве иллюстрации использования метода передискретизации показана реализация демодулятора сигнала с набором чиповых скоростей от 0,5 до 20 МГц. Демодулятор сочетает в себе децимирующие CIC-фильтры с коэффициентом децимации, равным степени двойки, и передискретизаторы с дробным рациональным коэффициентом преобразования частоты дискретизации.

Приводятся результаты компьютерного моделирования предложенного демодулятора. Качественный анализ результатов моделирования по «сжатию» точек фазовой диаграммы свидетельствует о малых энергетических потерях предлагаемого демодулятора. Следует особо подчеркнуть, что предложенный подход позволяет реализовать демодуляторы с непрерывной перестройкой по чиповой скорости. Это обеспечивает реализацию демодуляторов в различных системах навигации и связи. В частности, в системах спутниковой навигации и системах связи с кодовым разделением каналов. В заключение представляется целесообразным остановиться на обсуждении аппаратной реализации демодулятора. Во-первых, все представленные здесь результаты моделирования были получены с использованием арифметики с фиксированной точкой. Соответственно предложенный демодулятор не требует существенных вычислительных затрат и может быть реализован, в частности, на элементах программируемой логики. Во-вторых, реализация передискретизатора не требует значительных объемов памяти. ROM для хранения коэффициентов полиномов Лагранжа в рассматриваемой модели демодулятора составляет всего 16 кбайт. Во-третьих, выбор структуры с передискретизацией представляется предпочтительным в контексте общего построения цифрового приемника, так как предлагаемая реализация демодулятора не требует перестройки частоты дискретизации и фильтров, ограничивающих полосу сигнала на входе АЦП.

Список источников

1. Gardner F.M. Interpolation in digital modems. Part I: Fundamentals // IEEE Transactions on Communications. 1993. Vol. 41. Iss. 3. PP. 501–507. DOI:10.1109/26.221081

2. Erup L., Gardner F.M., Harris R.A. Interpolation in digital modems. Part II: Implementation and performance // IEEE Transactions on Communications. 1993. Vol. 41. Iss. 6. PP. 998–1008. DOI:10.1109/26.231921

3. ГЛОНАСС. Принципы построения и функционирования. Под ред. А.И. Перова, В.Н. Харисова. М.: Радиотехника, 2010. 800 с.

4. Кинкулькин И.Е. Глобальные навигационные спутниковые системы. Алгоритмы функционирования аппаратуры потребителя. М.: Изд-во «Едитория УРСС», 2018. 325 с.

5. Rec. ITU-R TF.1153-4 (08/2015). The operation use of two-way satellite time and frequency transfer employing pseudorandom noise code.

6. Gardner F.M. Phaselock Techniques. John Wiley & Sons, 2005. 450 p.

7. Mengali U., D'Andrea A.N. Synchronization Technique for Digital Receivers. New York: Plenum Press, 1997.

8. Meyer H., Moeneclaey M., Fechtel S.A.H. Digital Communication Receivers. New York: John Wiley & Sons, 1998.

9. Farrow C.W. A continuously variable digital delay element // Proceedings of the IEEE International Symposium on Circuits and Systems (Espoo, Finland, 7–9 June 1988). IEEE, 1988. PP. 2641–2645. DOI:10.1109/ISCAS.1988.15483

10. Hogenauer E. An economical class of digital filters for decimation and interpolation // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1981. Vol. 29. Iss. 2. PP. 155–162. DOI:10.1109/TASSP.1981.1163535

11. Брусин Е.А. Реализация начальной синхронизации демодулятора сигнала с прямым расширением спектра с использованием частотной автоподстройки // XIII Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (Санкт-Петербург, Российская Федерация, 27–28 февраля 2024 г.). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2024. С. 504–509. EDN:ZGFNZS

References

1. Gardner F.M. Interpolation in digital modems. Part I: Fundamentals. *IEEE Transactions on Communications*. 1993;41(3): 501–507. DOI:10.1109/26.221081
2. Erup L., Gardner F.M., Harris R.A. Interpolation in digital modems. Part II: Implementation and performance. *IEEE Transactions on Communications*. 1993;41(6):998–1008. DOI:10.1109/26.231921
3. *GLONASS. Principles of Construction and Functioning.* Edited by A.I. Perov, V.N. Kharisov. Moscow: Radiotekhnika Publ.; 2010. 800 p. (in Russ.)
4. Kinkulkin I.E. *Global Navigation Satellite Systems. Functioning Algorithms of Consumer Equipment*. Moscow: Editoriia URSS Publ.; 2018. 325 p. (in Russ.)
5. Rec. ITU-R TF.1153-4. *The operation use of two-way satellite time and frequency transfer employing pseudorandom noise code*. August 2015.
6. Gardner F.M. *Phaselock Techniques*. John Wiley & Sons; 2005. 450 p.
7. Mengali U., D'Andrea A.N. *Synchronization Technique for Digital Receivers.* New York: Plenum Press; 1997.
8. Meyer H., Moeneclaey M., Fechtel S.A.H. Digital Communication Receivers. New York: John Wiley & Sons; 1998
9. Farrow C.W. A continuously variable digital delay element. *Proceedings of the IEEE International Symposium on Circuits and Systems, 7–9 June 1988, Espoo, Finland*. IEEE; 1988. p.2641–2645. DOI:10.1109/ISCAS.1988.15483
10. Hogenauer E. An economical class of digital filters for decimation and interpolation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1981;29(2):155–162. DOI:10.1109/TASSP.1981.1163535
11. Brusin E. Implementation Direct Spread Spectrum Signals Demodulator Acquisition Using Automatic Frequency Control. *Proceedings of the XIIIth International Conference on Infotelecommunications in Science and Education, 27–28 February 2024, St. Petersburg, Russian Federation*. St. Petersburg: The Bonch-Bruevich Saint-Petersburg State University of Telecommunications Publ.; 2024. p.504–509. (in Russ.) EDN:ZGFNZS

Статья поступила в редакцию 28.10.2024; одобрена после рецензирования 09.12.2024; принята к публикации 11.12.2024.
The article was submitted 28.10.2024; approved after reviewing 09.12.2024; accepted for publication 11.12.2024.

Информация об авторе:

БРУСИН Ефим Александрович

кандидат технических наук, руководитель проекта Института радионавигации и времени АО «Обуховский завод», доцент кафедры Электроники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0002-6742-2705

Автор сообщает об отсутствии конфликтов интересов.
The author declares no conflicts of interests.

Научная статья
УДК 621.382
https://doi.org/10.31854/1813-324X-2024-10-6-19-25

Оценка возможности формирования канала утечки информации из оптического волокна тепловым воздействием

Иван Романович Гулаков, gulakov@bsu.by
Андрей Олегович Зеневич, a.zenevich@bsac.by
Татьяна Александровна Матковская ✉, tandem7m@gmail.com
Евгений Владимирович Новиков, e.novikov@bsac.by

Белорусская государственная академия связи,
Минск, 220114, Республика Беларусь

Аннотация

Работа посвящена оценке возможности формирования канала утечки информации с дефекта оптического волокна, созданного путем теплового воздействия. Свойства неоднородностей оптического волокна, вызванные таким воздействием, на сегодняшний день практически не изучены, что определяет ***актуальность*** *исследований. С учетом вышесказанного,* ***целью исследования*** *является определение характеристик неоднородностей оптического волокна, вызванных тепловым воздействием.*

Используемые методы*. В работе проведен расчет потерь мощности излучения, вносимых дефектом, вызванным тепловым воздействием при высокой температуре, а также мощности излучения, отводимой с дефекта за пределы оптического волокна. В ходе исследований характеристики неоднородностей оптического волокна, вызванных тепловым воздействием, оценивались также и по рефлектограммам.*

Результат*. В работе показано, что при помощи локального температурного воздействия удается сформировать дефект оптического волокна, позволяющий выводить часть оптического излучения за пределы этого волокна, то есть создать канал несанкционированного съема данных. Величина вносимых потерь мощности излучения на создаваемом дефекте возрастала с увеличением времени теплового воздействия на оптическое волокно. При времени теплового воздействия на оптическое волокно менее 1 с сформировать дефект с существенными вносимыми потерями мощности излучения не удавалось, а при времени теплового воздействия более 10 с вносимые потери на дефекте превышали 20 дБ (в этом случае прекращается передача данных зональных и магистральных ВОЛС). Показано, что с увеличением длины волны распространяющегося по волокну оптического излучения возрастают потери мощности излучения на дефекте, сформированном тепловым воздействием на оптическое волокно. Установлено, что при одинаковой потере мощности на дефекте, сформированном тепловым воздействием, мощность оптического излучения, отводимая с такого дефекта, имеет наибольшее значение при использовании оптического волокна G652, а наименьшее – при использовании волокна G657.*

Научная новизна *работы состоит в исследовании ранее неизученных свойств неоднородностей оптического волокна, вызванных тепловым воздействием.*

Практическая значимость*. Результаты, приведенные в статье, могут найти применение при проектировании систем защиты информации, передаваемой по волоконно-оптическим линиям связи*

Ключевые слова: *оптическое волокно, дефект оптического волокна, тепловое воздействие, канал утечки информации*

Ссылка для цитирования: Гулаков И.Р., Зеневич А.О., Матковская Т.А., Новиков Е.В. Оценка возможности формирования канала утечки информации из оптического волокна тепловым воздействием // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 19–25. DOI:10.31854/1813-324X-2024-10-6-19-25. EDN:MQXMKY

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-19-25

Assessment of Forming Information Leakage Channel from Optical Fiber Possibility by Thermal Exposure

Ivan R. Gulakov, gulakov@bsu.by
Andrey O. Zenevich, a.zenevich@bsac.by
Tatyana A. Matkovskaia ✉, tandem7m@gmail.com
Evgeniy V. Novikov, e.novikov@bsac.by

Belarusian State Academy of Communications,
Minsk, 220114, Republic of Belarus

Annotation

The article is devoted to assessing the possibility of forming an information leakage channel from an optical fiber defect created by thermal exposure. The properties of optical fiber inhomogeneities caused by such exposure have not been practically studied to date, which determines the ***relevance*** *of research. Taking into account the above, the* ***purpose of the study*** *is to determine the characteristics of optical fiber inhomogeneities caused by thermal exposure.*
The methods used. *The paper calculates the radiation power losses introduced by a defect caused by thermal action at high temperature, as well as the radiation power removed from the defect beyond the optical fiber. During the studies, the characteristics of optical fiber inhomogeneities caused by thermal action were also estimated using reflectograms.*
The result. *The work that local temperature exposure makes it possible to form a defect in an optical fiber that allows part of the optical radiation to be emitted beyond the fiber, i.e. to create a channel for unauthorized data retrieval. The magnitude of the insertion loss of radiation power on the created defect increased with increasing time of thermal exposure to the optical fiber. When the time of thermal exposure to the optical fiber was less than 1 s, it was not possible to form a defect with significant insertion loss of radiation power, and when the time of thermal exposure was more than 10 s, the insertion loss on the defect exceeded 20 dB, at which data transmission of zonal and trunk fiber-optic communication lines ceases. It is shown that with increasing wavelength of optical radiation propagating along the fiber, the loss of radiation power on the defect formed by thermal exposure to the optical fiber increases. It has been established that with the same power loss on a defect formed by thermal action, the optical radiation power removed from such a defect has the greatest value when using G652 optical fiber, and the least when using G657 fiber.*
The scientific novelty *of the work consists in the study of previously unexplored properties of optical fiber inhomogeneities caused by thermal exposure.*
Practical Significance. *The results presented in the article can be used in the design of information protection systems transmitted over fiber-optic communication lines.*

Keywords: *optical fiber, optical fiber defect, thermal effect, information leakage channel*

For citation: Gulakov I.R., Zenevich A.O., Matkovskaia T.A., Novikov E.V. Assessment of Forming Information Leakage Channel from Optical Fiber Possibility by Thermal Exposure. *Proceedings of Telecommunication Universities.* 2024;10(6):19–25. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-19-25. EDN:MQXMKY

Введение

На сегодняшний день для передачи информации активно применяются волоконно-оптические линии связи (ВОЛС). Оптические волокна, входящие в состав ВОЛС, имеют преимущества в скорости и пропускной способности по сравнению с медными жилами [1–4]. Данные, передаваемые по оптическим волокнам, более защищены от несанкционированного доступа к передаваемой информации, однако возможно формирование канала утечки информации путем отвода части оптического излучения из волокна без его разрыва [5]. К способам реализации канала утечки информации относятся формирование макроизгиба и микроизгиба оптического волокна [6–9], а также способы, реализованные на основе оптического туннелирования и сдавливания оптического волокна [10].

В случае подключения устройств, реализующих вышеуказанные способы, обычно требуется доступ к отрезку оптического волокна определенной протяженности, однако не всегда удается его получить. Вместе с тем оказывается возможным создать зону подключения устройства, обеспечивающего несанкционированный съем информации, путем теплового воздействия на волокно. Такое воздействие приводит к появлению его неоднородности, с которой возможен выход части мощности оптического излучения за пределы волокна.

Для обнаружения неоднородностей оптического волокна, вызванных тепловым воздействием при высокой температуре, необходимо знать характеристики этих неоднородностей. Наиболее важными характеристиками неоднородностей являются вносимые ими потери мощности излучения в оптическом волокне и доля мощности оптического излучения, отводимая через сформированную неоднородность за пределы оптического волокна. Первое из этих свойств определяет возможность обнаружения наличия неоднородности. Второе позволяет оценить возможность применения неоднородности для съема данных. Однако свойства неоднородностей оптического волокна, вызванные тепловым воздействием, на сегодняшний день практически не исследованы. Поэтому целью настоящей работы является определение характеристик неоднородностей оптического волокна, вызванных тепловым воздействием.

Экспериментальная установка и методика измерения

Объектами исследований выбраны стандартные одномодовые оптические волокна G652, G655 и G657, достаточно часто применяемые в оптических кабелях. Свойства неоднородностей оптического волокна, вызванных тепловым воздействием, проанализированы на экспериментальной установке, структурная схема которой приведена на рисунке 1.

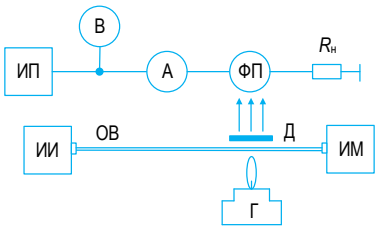


Рис. 1. Структурная схема экспериментальной установки

Fig. 1. Block Diagram of the Experimental Setup

В схеме приняты обозначения: ИИ – источник излучения; ОВ – оптическое волокно; ИМ – измеритель мощности; Д – диафрагма; ФП – фотоприемник; Г – горелка; А – амперметр; В – вольтметр; ИП – источник питания; $R_н$ – резистор нагрузки.

Принцип работы экспериментальной установки: от источника ИИ оптическое излучение поступает в волокно ОВ, а из него – на измеритель мощности ИМ. От источника постоянного напряжения ИП на фотоприемник ФП подается напряжение питания $U_п$. Вольтметр В необходим для контроля величины напряжения питания. Амперметр А используется для измерения электрического тока, протекающего через фотоприемник ФП. Последовательно с фотоприемником включен резистор нагрузки $R_н$ номиналом 1 кОм, необходимый для ограничения величины тока, протекающего через фотоприемник.

Длина волны оптического излучения ИИ в процессе измерений могла изменяться и принимать значения 1310, 1490, 1550 и 1650 нм, соответствующие «окнам прозрачности» одномодовых оптических волокон. В ходе проведения исследования не учитывались потери мощности оптического излучения в волокне, так как его длина составляла всего 1 м. Отметим, что для всех используемых оптических волокон на исследуемых длинах волн потеря мощности излучения не превышает 0,4 дБ/км.

Для формирования дефекта в оптическом волокне ОВ небольшая часть этого волокна помещалось в пламя горелки Г. При этом диафрагма Д закрывалась, чтобы ограничивать попадания света и тепла от горелки Г на фотоприемник ФП.

При проведении исследований вычислялся фототок $I_ф$ фотоприемника ФП:

$$Iф = I - Iт, \quad (1)$$

где $I_т$ – темновой электрический ток, измеряемый при закрытой диафрагме Д; I – электрический ток, измеряемый при открытой диафрагме.

По величине фототока определялась мощность оптического излучения, поступающего на фотоприемник ФП с неоднородности оптического волокна, вызванной температурным воздействием: $P_{отв} = I_ф/S$, где S – чувствительность фотоприемника.

Отметим, что различные области пламени имеют различную температуру [11]. Так, температура верхней области пламени наибольшая (1500 К), а температура области, находящаяся возле фитиля – наименьшая (800 К) [12]. Поэтому оптическое волокно помещалось в верхнюю часть пламени. Это позволяло подвергать волокно тепловому воздействию с постоянной температурой 1500 К.

При выполнении измерений потери мощности излучения, вносимые дефектом, вызванным температурным воздействием, определялись по следующей формуле:

$$D_п = 10\lg\left(\frac{P}{P_в}\right), \quad (2)$$

где P – мощность источника излучения ИИ; $P_в$ – мощность оптического излучения, поступающая на измеритель мощности ИМ.

Время теплового воздействия изменялось от 1 до 10 с. Чем более длительное воздействие осуществлялось на оптическое волокно, тем бóльшими были вносимые потери мощности излучения на создаваемом дефекте. Отметим, что при времени теплового воздействия на оптическое волокно менее 1 с сформировать дефект с существенными вносимыми потерями мощности излучения $D_{\text{п}}$ не удавалось. При времени теплового воздействия на оптическое волокно более 10 с вносимые потери мощности излучения $D_{\text{п}}$ на дефекте превышали 20 дБ. Такая величина $D_{\text{п}}$ для зональных и магистральных ВОЛС приводит к прекращению передачи данных. Поэтому рассматривать воздействия, вносящие такие потери мощности излучения, считалось нецелесообразным.

Длина участка, на котором возникал вызванный тепловым воздействием дефект, составляла около 1 см. Этого было достаточно для регистрации величины мощности, отводимой за пределы оптического волокна с неоднородности, вызванной тепловым воздействием. Отметим, к примеру, что несанкционированное снятие данных с оптического волокна при помощи микроизгиба требует доступа к участку оптического волокна протяженностью от 1,5 до 3 см, а в случае использования макроизгиба – от 3 до 5 см.

Под влиянием температуры в области воздействия на оптическое волокно сгорало лакокрасочное покрытие волокна, если такое имелось. После теплового воздействия в области сформированного дефекта волокно становилось хрупким и даже при незначительном изгибе ломалось. В отсутствии лакокрасочного покрытия внешний вид области, на которую осуществлялось температурное воздействие, оставался таким же, как и до него.

Для определения мощности излучения $D_{\text{отв}}$, отводимой за пределы оптического волокна с неоднородности, вызванной тепловым воздействием, применялось следующее выражение:

$$D_{\text{отв}} = 10\lg\left(\frac{P_{\text{отв}}}{P}\right). \qquad (3)$$

В ходе исследований характеристики неоднородностей оптического волокна, вызванных тепловым воздействием, оценивались также и по рефлектограммам, полученным на экспериментальной установке, структура которой приведена в работе [13]. В этом случае неоднородность, вызванная тепловым воздействием, формировалась в середине оптического волокна протяженностью 1,5 км. При проведении измерений регистрировались рефлектограммы сигналов в волокне с наличием такой неоднородности, для чего ко входу оптического волокна подключался рефлектометр. Длительность оптического импульса рефлектометра составляла 3 нс. При такой длительности оптических импульсов длина мертвой зоны по затуханию используемого рефлектометра была минимальной, что позволяло с наибольшей точностью определять местоположения теплового воздействия и величину потерь мощности излучения на сформированном дефекте.

Измерения величин $D_{\text{п}}$ и $D_{\text{отв}}$, а также рефлектограмм выполнялись при комнатной температуре T = 293 К.

Результаты измерений и их обсуждения

В процессе исследований создавался дефект оптического волокна путем теплового воздействия на оптическое волокно пламенем горелки (см. рисунок 1). При этом величина потери мощности на дефекте зависела от времени воздействия на оптическое волокно пламени, подвергающегося воздействию пламени.

При тепловом воздействии в оптическом волокне может происходить диффузия примесей, введенных в это волокно при его производстве. Такая диффузия приводит к изменению абсолютных значений показателей преломления сердцевины и оболочки оптического волокна в месте температурного воздействия, а, следовательно, и их разности. Это приводит к нарушению условия существования одной моды в одномодовом волокне и появлению дополнительных мод. Перераспределение энергии между модами приводит к потере мощности передаваемого оптического сигнала и выходу энергии дополнительных мод за пределы волокна в области его локального нагрева. Выходу энергии оптического излучения за пределы волокна способствует также удаление лакокрасочного покрытия.

Результаты измерений потери мощности оптического излучения для одного и того же дефекта, вызванного тепловым воздействием на оптическое волокно, для разных длин волн оптического излучения, приведены в таблице 1. Как следует из полученных результатов, чем больше длина волны, тем больше потери мощности оптического излучения для всех исследуемых оптических волокон. В таблице 1 также отображены сведения об отведении мощности оптического излучения с боковой поверхности оптического волокна в области дефекта. Как видно из полученных данных, с увеличением длины волны отводится большая мощность оптического излучения с боковой поверхности оптического волокна в области дефекта. Вышеописанные тенденции наблюдаются для всех исследуемых оптических волокон. Отметим, что дефекты оптических волокон G652, обладающих меньшей потерей мощности оптического излучения, чем другие исследуемые оптические волокна, имели большие

значения отводимой мощности оптического излучения с боковой поверхности оптического волокна в области дефекта. Это связано с различной внутренней структурой исследуемых волокон.

ТАБЛИЦА 1. Результаты измерений потери мощности оптического излучения в области дефекта, вызванного тепловым воздействием на оптическое волокно

TABLE 1. Measurements of Optical Radiation Power Loss in the Area of Defect Caused by Thermal Effects on the Optical Fiber Results

Тип ОВ	λ, нм	Потери мощности оптического излучения, дБ	Отведение мощности оптического излучения, дБ
G652	1310	0,40	–48,70
	1490	1,23	–48,10
	1550	1,29	–47,40
	1625	1,51	–46,90
G655	1310	0,97	–56,60
	1490	2,16	–56,10
	1550	2,79	–55,60
	1625	3,52	–55,10
G657	1310	5,01	–57,80
	1490	5,27	–57,20
	1550	5,49	–56,30
	1625	6,10	–54,30

Анализ участка рефлектограммы оптического волокна G652, содержащего дефект, показывает, что для места нахождения такого дефекта характерно наличие перепада мощности в виде «ступеньки» (рисунок 2). Величина такого перепада мощности увеличивается с ростом длины волны, т. е. для этих дефектов наблюдается зависимость потери мощности оптического излучения от длины волны оптического излучения. Отметим, что рефлектограммы для этих дефектов и их поведение с ростом длины волны оптического излучения схожи с рефлектограммами, характерными для макроизгибов оптического волокна [14].

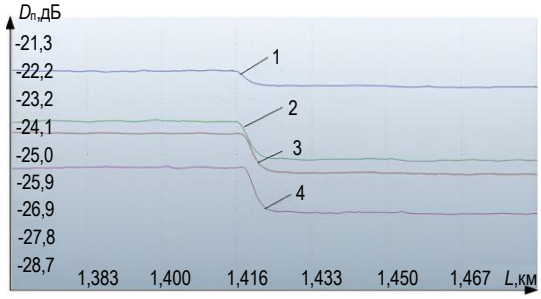


Рис. 2. Участок рефлектограммы оптического волокна G652, содержащего дефект, для длин волн: 1 – 1310 нм, 2 – 1490 нм, 3 – 1550 нм, 4 – 1625 нм

Fig. 2. The Section of Optical Fiber G652 Reflectogram Containing Defect for Wavelengths: 1 – 1310 nm, 2 – 1490 nm, 3 – 1550 nm, 4 – 1625 nm

Рефлектограммы приведены для оптического волокна G652. Для других оптических волокон рефлектограммы идентичны. На рефлектограммах можно увидеть, что увеличение потери мощности оптического излучения на дефекте приводит к росту величины мощности оптического излучения, отводимой с дефекта. При одинаковом значении потери мощности на дефекте для разных длин волн оптического излучения наблюдалась различная величина отводимой с дефекта мощности оптического излучения.

На рисунке 3 представлены типичные зависимости величины мощности излучения $P_{отв}$, отводимой с дефекта, сформированного в результате теплового воздействия, от величины потери мощности излучения, возникшей из-за этого дефекта, для различных длин волн. Как следует из полученных зависимостей, с увеличением потери мощности излучения растет значение мощности излучения, отводимой с дефекта $P_{отв}$. Полученные зависимости имели нелинейный вид. Это свидетельствует о том, что потери мощности на таком дефекте обусловлены не только излучением, выходящим за пределы оптического волокна в области этого дефекта.

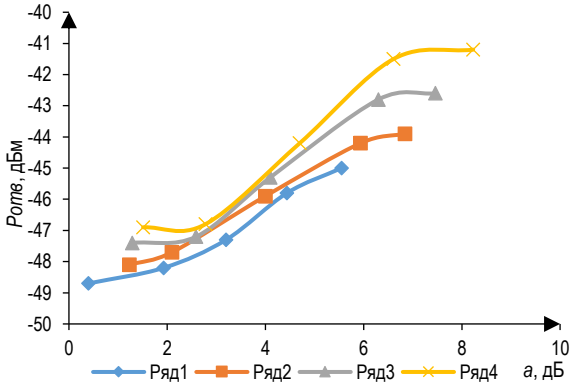


Рис. 3. Зависимость отведения мощности оптического излучения с дефекта, сформированного в результате теплового воздействия, от величины потери мощности, возникшей из-за этого воздействия, для длин волн: 1 – 1310 нм, 2 – 1490 нм, 3 – 1550 нм, 4 – 1625 нм

Fig. 3. The Dependence of Optical Radiation Power Removal from Defect Formed as Thermal Exposure Result on Power Loss Magnitude Caused by this Effect for Wavelengths: 1 – 1310 nm, 2 – 1490 nm, 3 – 1550 nm, 4 – 1625 nm

В процессе проведенного исследования было выполнено сравнение отводимой с дефекта мощности оптического излучения для разных оптических волокон. При проведении сравнения в каждом из исследуемых оптических волокон создавались дефекты, которые вносили одинаковую потерю мощности на одной и той же длине волны оптического излучения. Для этой же длины волны оптического излучения измерялась потеря мощности оптического излучения. Сведения о полученных результатах представлены в таблице 2. Исходя из представленных данных, наибольшее значение мощности, отводимой с сформированного в результате теплового воздействия дефекта, наблюдается для оптического волокна G652, а наименьшее – для G657.

Это наблюдается для всех исследуемых длин волн излучения. Такое отличие в значениях отводимой мощности с дефекта связано с различной внутренней структурой исследуемых оптических волокон. Отметим, что исследуемые волокна имеют отличие в геометрических размерах сердцевины и оболочки волокна [15–17].

ТАБЛИЦА 2. Результаты измерений отведения мощности оптического излучения с боковой поверхности оптического волокна в области дефекта, сформированного тепловым воздействием

TABLE 2. Measurements of Optical Radiation Power Removal from Lateral Surface of Optical Fiber in the Defect Area Formed by Thermal Action Results

Тип ОВ	Длина волны, нм	Потеря мощности, дБ	Отведение мощности с боковой поверхности волокна в области дефекта, дБ
G652	1310	3,5	–47,1
G655			–52,8
G657			–58,6
G652	1490	4,5	–46,0
G655			–52,0
G657			–58,0
G652	1550	5,0	–44,1
G655			–51,0
G657			–56,5
G652	1625	6,0	–42,0
G655			–50,7
G657			–54,4

Сравнение величины мощности, отводимой за пределы оптического волокна с неоднородности, вызванной тепловым воздействием, со значениями мощностей, отводимых из волокна в области специально сформированных микро- или макроизгибов, показало, что они составляли –46, –35 и –20 дБ, соответственно, для теплового воздействия, микро- и макроизгибов. Сравнение выполнялось для оптического волокна G652 и длины волны 1310 нм в условиях одинаковой потери мощности в 4 дБ на каждом из видов неоднородностей. Наименьшая по величине мощность отводится за пределы оптического волокна с неоднородности, вызванной тепловым воздействием, однако такого значения мощности достаточно, чтобы обеспечить несанкционированный съем данных с этой неоднородности.

Заключение

Показано, что при помощи локального температурного воздействия удается сформировать дефект оптического волокна, позволяющий выводить часть оптического излучения за пределы этого волокна. Определено, что при увеличении длины волны возрастает мощность оптического излучения, отводимого с дефекта, сформированного тепловым воздействием на оптическое волокно. Исследования показали, что при одинаковой потере мощности на дефекте, сформированном тепловым воздействием на оптическое волокно, отводимая мощность оптического излучения с такого дефекта имеет наибольшее значение при использовании оптического волокна G652, а наименьшее – при использовании G657.

Список источников

1. Скляров О.К. Волоконно-оптические сети и системы связи. СПб.: Лань, 2021. 268 с.
2. Senior J.M., Jamro M.Y. Optical fiber communications: principles and practice. Financial Times/Prentice Hall, 2009. 1127 p.
3. Ионов А.Д. Волоконно-оптические линии передачи. Новосибирск: СибГУТИ, 2003. 152 с.
4. Govind P. Agrawal Fiber-Optic Communication Systems. New York: Wiley-Interscience, 2002. 563 p.
5. Зеневич А.О. Обнаружители утечки информации из оптического волокна. Минск: Белорусская государственная академия связи, 2017. 142 с.
6. Гулаков И.Р., Зеневич А.О., Кочергина О.В., Матковская Т.А. Исследование канала утечки информации в области изгиба оптического волокна // Труды учебных заведений связи. 2022. Т. 8. № 3. С. 44–49. DOI:10.31854/1813-324X-2022-8-36-44-49. EDN:CPHMYU
7. Гулаков И.Р., Зеневич А.О, Матковская Т.А., Новиков Е.В. Исследования свойств микроизгиба одномодового оптического волокна // Труды учебных заведений связи. 2023. Т. 9. № 4. С. 15–20. DOI:10.31854/1813-324X-2023-9-4-15-20. EDN:QFEQEF
8. Wang Q., Farrell G., Freir T. Theoretical and Experimental Investigations of Macro-Bend Losses for Standard Single Mode Fibers // Optics Express. 2005. Vol. 13. Iss. 12. PP. 4476–4484. DOI:10.1364/OPEX.13.004476
9. Schermer R.T., Cole J.H. Improved Bend Loss Formula Verified for Optical Fiber by Simulation and Experiment // IEEE Journal of Quantum Electronics. 2007. Vol. 43. Iss. 10. PP. 899–909. DOI:10.1109/JQE.2007.903364
10. Шубин В.В. Информационная безопасность волоконно-оптических систем. Саров: РФЯЦ-ВНИИЭФ, 2015. 257 с.
11. Девисилов В.А., Дроздова Т.И., Тимофеева С.С. Теория горения и взрыва: учебное пособие. М.: ФОРУМ, 2012. 352 с.
12. Стариков А.Н. Основы теории горения и взрыва. Владимир: Изд-во ВлГУ, 2019. 148 с.
13. Листвин А.В., Листвин В.Н. Рефлектометрия оптических волокон. М.: ЛЕСАРарт, 2005. 208 с.
14. Зеневич А.О., Новиков Е.В., Матковская Т.А., Горбадей О.Ю., Василевский Г.В. Обнаружение изгибов оптического волокна вблизи сварных и механических соединений // Проблемы инфокоммуникаций. 2022. № 2(16). С. 32–38. EDN:QGIWAB
15. Рекомендация МСЭ-Т G652 (11/2016). Характеристики одномодового оптического волокна и кабеля.
16. Рекомендация МСЭ-Т G655 (11/2009). Характеристики одномодового оптического волокна и кабеля с ненулевой смещенной дисперсией.
17. Рекомендация МСЭ-Т G657 (11/2016). Характеристики одномодового оптического волокна и кабеля, не чувствительного к потерям на изгибе.

References

1. Sklyarov O.K. *Fiber-optic networks and communication systems*. St. Petersburg: Lan Publ.; 2021. 268 p. (in Russ.)
2. Senior J.M., Jamro M.Y. *Optical fiber communications: principles and practice*. Financial Times/Prentice Hall; 2009. 1127 p.
3. Ionov A.D. *Fiber-Optic Transmission Lines*. Novosibirsk: SibSUTI Publ.; 2003. 152 p. (in Russ.)
4. Govind P. *Agrawal Fiber-Optic Communication Systems*. New York: Wiley-Interscience; 2002. 563 p.
5. Zenevich A.O. *Optical Fiber Information Leak Detectors.* Minsk: Belarusian State Academy of Communications Publ.; 2017. 142 p. (in Russ.)
6. Gulakov I., Zenevich A., Kochergina O., Matkovskaia T. Investigation of an Information Leakage Channel in the Area Optical Fiber Bending. *Proceedings of Telecommunication Universities.* 2022;8(3):44–49. (in Russ.) DOI:10.31854/1813-324X-2022-8-36-44-49. EDN:CPHMYU
7. Gulakov I., Zenevich A., Matkovskaya T., Novikov E. Investigations of Single-Mode Optical Fiber Microbending Properties. *Proceedings of Telecommunication Universities*. 2023;9(4):15–20. (in Russ.) DOI:10.31854/1813-324X-2023-9-4-15-20. EDN:QFEQEF
8. Wang Q., Farrell G., Freir T. Theoretical and Experimental Investigations of Macro-Bend Losses for Standard Single Mode Fibers. *Optics Express*. 2005;13(12):4476–4484. DOI:10.1364/OPEX.13.004476.
9. Schermer R.T., Cole J.H. Improved Bend Loss Formula Verified for Optical Fiber by Simulation and Experiment. *IEEE Journal of Quantum Electronics*. 2007;43(10):899–909. DOI:10.1109/JQE.2007.903364
10. Shubin V.V. *Information Security of Fiber-Optic Systems*. Sarov: RFNC-VNIIEF Publ.; 2015. 257 p. (in Russ.)
11. Devisilov V.A., Drozdova T.I., Timofeeva S.S. *Theory of Combustion and Explosion.* Moscow: FORUM Publ.; 2012. 352 p. (in Russ.)
12. Starikov A.N. *Fundamentals of the Theory of Combustion and Explosion*. Vladimir: VlSU Publ.; 2019. 148 p. (in Russ.)
13. Listvin A.V., Listvin V.N. *Reflectometry of Optical Fibers*. Moscow: LESARart Publ.; 2005. 208 p. (in Russ.)
14. Zenevich A.O., Novikov E.V., Matkovskaya T.A., Gorbaday O.Yu., Vasilevsky G.V. Detection of Bends of Optical Fiber Near Welded And Mechanical Joints. *Problems of infocommunications*. 2022;2(16):32–38. (in Russ.) EDN:QGIWAB
15. Rec. ITU-T G652. *Characteristics of a single-mode optical fibre and cable*. 2016.
16. Rec. ITU-T G655. *Characteristics of a non-zero dispersion-shifted single-mode optical fibre and cable*. November 2009.
17. Rec. ITU-T G657. *Characteristics of a bending-loss insensitive single-mode optical fibre and cable*. November 2016.

Статья поступила в редакцию 15.09.2024; одобрена после рецензирования 25.11.2024; принята к публикации 05.12.2024.
The article was submitted 15.09.2024; approved after reviewing 25.11.2024; accepted for publication 05.12.2024.

Информация об авторах:

ГУЛАКОВ Иван Романович — доктор физико-математических наук, профессор, профессор кафедры физических и математических основ информатики Белорусской государственной академии связи
https://orcid.org/0000-0002-7330-9928

ЗЕНЕВИЧ Андрей Олегович — доктор технических наук, профессор, ректор Белорусской государственной академии связи
https://orcid.org/0000-0002-3534-3885

МАТКОВСКАЯ Татьяна Александровна — аспирант кафедры физических и математических основ информатики Белорусской государственной академии связи
https://orcid.org/0000-0002-1499-6158

НОВИКОВ Евгений Владимирович — кандидат технических наук, доцент, директор Института современных технологий связи Белорусской государственной академии связи
https://orcid.org/0009-0009-2944-758X

Зеневич А.О. является членом редакционного совета журнала «Труды учебных заведений связи» с 2023 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.

Zenevich A.O. has been a member of the journal "Proceedings of Telecommunication Universities" Editorial Council since 2023, but has nothing to do with the decision to publish this article. The article has passed the review procedure accepted in the journal. The authors have not declared any other conflicts of interest.

Научная статья
УДК 621.372.2
https://doi.org/10.31854/1813-324X-2024-10-6-26-33

Синтез устройств СВЧ диапазона на основе микроволнового кольцевого эллиптического резонатора

Александр Сергеевич Леонтьев ✉, leontev.as@sut.ru
Эрнест Юрьевич Седышев, k112_electron@sut.ru

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

Аннотация

Актуальность. *Исследования кольцевых резонирующих структур вызывает интерес у разработчиков СВЧ устройств, некоторые особенности кольцевых эллиптических резонаторов приводят к возникновению уникальных свойств передаточных характеристик. Особое значение при использовании данных резонаторов имеет способ их возбуждения. При определенных условиях в этих структурах возможно получение режима бегущей волны. Синтезу микроволновых устройств, а также исследованию способов возбуждения и оценки режима волнового процесса в структуре посвящена данная работа.*

Цель исследования *– проанализировать и систематизировать информацию о создании микроволновых устройств с использованием кольцевых эллиптических резонаторов (КЭР), а также апробировать результаты применения сдвоенного КЭР.*

Методы: *в этой работе был проведен аналитический обзор актуальных научных публикаций, а также выполнено компьютерное моделирование полосковых кольцевых эллиптических резонаторов, работающих в сверхвысокочастотном диапазоне. В работе также представлены результаты экспериментов, апробированные разными исследователями, в том числе поддержанные грантами РФФИ.*

Решение. *В статье рассматриваются особенности использования шлейфных полосковых фильтров и описываются ограничения, возникающие при использовании полосковых резонаторов. Представлена конструкция КЭР и предлагается его применение в качестве альтернативы полосковым резонаторам. Приводятся результаты многочисленных экспериментов по синтезу микроволновых устройств на основе КЭР, включая: одинарный; двойной резонаторы; преселективные фильтры; усилители и генераторы, построенные на основе кольца и активных двухполюсниках. Также рассматривается проблема коммутации резонатора с основной линией передачи. Приводятся результаты макетирования нескольких устройств, ограничивающих направление распространения электромагнитной волны в кольцевом резонаторе.*

Научная новизна. *Впервые представлена конструкция двойного эллиптического резонатора, выполненного на основе микрополосковой линии. В статье приводятся результаты эксперимента, которые демонстрируют достижение уровня режекции фильтра более 70 дБ исключительно за счет топологии резонатора. Обсуждается проблема выбора способа питания и обеспечения режима распространения волны в КЭР.*

Практическая значимость: *результаты, полученные в ходе работы, могут быть применены для создания резонатора бегущей волны на микрополосковой линии или в других планарных или объемных конфигурациях. Также результаты исследования служат основой для создания обобщенной теории синтеза кольцевых резонаторов микроволнового диапазона длин волн.*

Ключевые слова: *кольцевой эллитпический резонатор, резонанс, фильтр, добротность, бегущая волна, синтез, СВЧ*

Ссылка для цитирования: Леонтьев А.С., Седышев Э.Ю. Синтез устройств СВЧ диапазона на основе микроволнового кольцевого эллиптического резонатора // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 26–33. DOI:10.31854/1813-324X-2024-10-6-26-33. EDN:EAONVX

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-26-33

Synthesis of Microwave Devices Based on a Microwave Ring Elliptical Resonator

Alexander S. Leontev, leontev.as@sut.ru
Ernest Y. Sedyshev, k112_electron@sut.ru

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

Annotation

Actuality. *Research of ring resonating structures is interesting to developers of microwave devices; some features of ring elliptical resonators lead to the emergence of unique properties of the transmission characteristics of the device. Method of excitation for CER is particularly important. It is possible to obtain the traveling wave mode in these structures, under certain conditions. To the synthesis of microwave devices, as well as to the study of methods for exciting and evaluating the mode of the wave process in the structure this paper is devoted.*

Object. *The purpose of the study is to analyze and organize information about the development of microwave devices using circular elliptical resonators (CERs). Authors also want to test the results of using dual CERs.*

Methods. *The authors have conducted an analytical review of recent scientific publications and performed computer modeling of microstrip ring elliptical resonators that operate in the ultrahigh frequency range in this work. The paper also includes the results of our experiments, tested by various researchers, including those supported by grants from the Russian Foundation for Basic Research.*

Result. *The article explores the unique characteristics of loop strip filters and highlights the limitations of using strip resonators. It that describes the design of a ring elliptical resonator (CER), and suggests its potential as an alternative to microstrip resonators. The paper presents the results of numerous experiments on the development of microwave devices based on CERs. The results of numerous experiments of the synthesis of microwave devices based on CER and including: single; double resonators; preselective filters; amplifiers and generators based on a ring and active bipolar are presented. Additionally, the issue of connecting the resonator to the main transmission line is addressed. The results of modeling several devices that limit the direction of propagation of an electromagnetic wave in an annular resonator are presented.*

Scientific novelty. *This article introduces a new design of a double elliptical resonator based on a microstrip line. It also describes the results of an experiment that shows that the resonator topology can achieve a filter rejection level of over 70 dB. In the article, the authors also discuss the problem of selecting a power supply method and ensuring the wave propagation mode in the CER.*

Practical significance. *The results obtained in the course of this work can be used to create a traveling wave resonator on a microstrip line or in other planar or volumetric configurations. The results of the study also serve as the basis for the creation of a generalized theory of synthesis of ring resonators in the microwave wavelength range.*

Keywords: *circular elliptical resonator, resonance, filter, Q-factor, traveling wave, synthesis, microwave*

For citation: Leontev A.S., Sedyshev E.Y. Synthesis of Microwave Devices Based on a Microwave Ring Elliptical Resonator. *Proceedings of Telecommunication Universities.* 2024;10(6):26–33. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-26-33. EDN:EAONVX

Введение

Постоянное развитие устройств и технологий СВЧ диапазона требует повышения степени интеграции основных топологических решений, учитывающих распространение волны в диэлектриках. Это неудивительно, так как с освоением более высоких частот все большее внимание уделяется конструктивным компонентам интегральных схем (ИС): соединительные линии, распределенные емкости, индуктивности, печатные фильтры, резонаторы и т. д.

Основная сложность в использовании подобных конструктивных элементов заключается в необходимости разработать их топологию и интегрировать в единый технологический цикл создания токонесущих слоев или напыления полупроводников и диэлектриков. На этом этапе также важно обеспечить возможность их настройки и корректировки.

В связи с этим разработка новых эффективных топологических решений приобретает особую значимость, поскольку традиционные подходы часто не могут удовлетворить противоречивым современным требованиям и не могут быть воспроизведены при переходе в более высокую часть СВЧ диапазона.

Особенности шлейфных фильтров

Основным конструктивным элементом, используемым при создании микроволновых устройств, является линия. Отрезок этой линии, называемый шлейфом, также служит основой для многих СВЧ устройств. Шлейфы позволяют создавать направленные ответвители, фильтры различных типов, а также осуществлять коммутацию и развязку микроволновых трактов. Кроме того, шлейфы играют важную роль как основные резонансные элементы в ИС сверхвысокой частоты. Однако микрополосковый шлейф имеет низкую добротность, что существенно ограничивает его применение. Для достижения требуемых характеристик фильтров, например, необходимо создавать структуры более высокого порядка.

Сравним графики *S*-параметров двух фильтров с Чебышевской характеристикой третьего и пятого порядков (рисунок 1). Фильтры были рассчитаны и смоделированы в программе RFSimm с центральной рабочей частотой 3 ГГц и полосой пропускания 100 МГц. Как видно из рисунков, фильтр более высокого порядка демонстрирует более высокую крутизну. Это означает, что он обладает лучшей избирательной способностью. Таким образом, для достижения заданной крутизны фильтра разработчик должен либо увеличить добротность резонаторов, выбрав соответствующий тип линии и материалы, либо повысить порядок самого фильтра. Чтобы оценить, насколько избирателен резонатор по частоте, используют величину, называемую добротностью.

Добротность показывает, как сильно отличается центральная частота от полосы резонанса, а также – сколько энергии запасает фильтр и сколько теряет за один период колебания:

$$Q = \frac{f_0}{\Delta f}, \quad (1)$$

где

$$\Delta f = f_2 - f_1 . \quad (2)$$

Характеристики рассчитанных фильтров приведены в таблице 1.

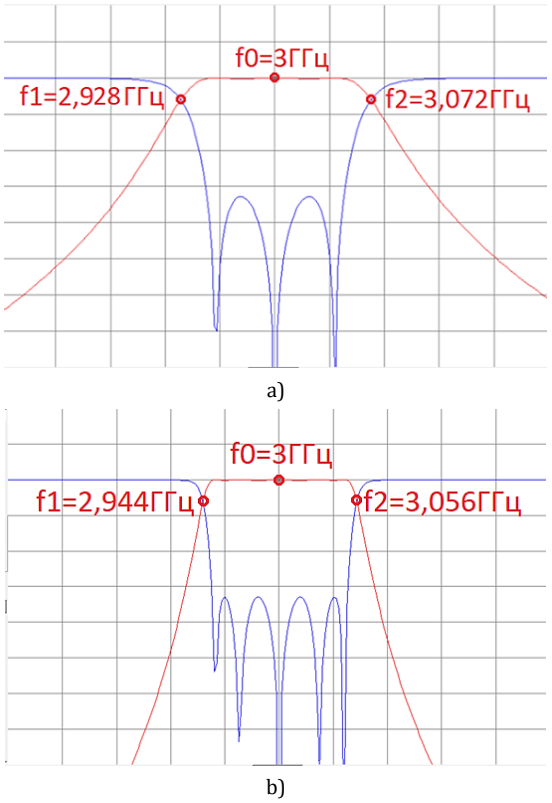


Рис. 1. *S*-параметры полоскового фильтров 3-го (a) и 5-го (b) порядков

Fig. 1. S-Parameters of 3rd (a) and 5th (b) Order Bandpass Filter

ТАБЛИЦА 1. Сравнение характеристик фильтров 3-го и 5-го порядков

TABLE 1. Characteristic Comparison of 3rd (a) and 5th (b) Order Bandpass Filter

Характеристики	Фильтры	
	3-го порядка	5-го порядка
Центральная частота	3 ГГц	
Частота нижнего среза	2,928 ГГц	2,944 ГГц
Частота верхнего среза	3,072 ГГц	3,056 ГГц
Полоса пропускания	144 МГц	112 МГц
Добротность	41,67	53,57

Таким образом, увеличение порядка фильтра можно рассматривать как повышение добротности резонаторов. В теории фильтрации обычно рассматриваются фильтры на линиях без потерь, и поэтому может возникнуть заблуждение, что увеличение порядка фильтра позволяет получить любую крутизну. Однако на самом деле крутизна фильтра в первую очередь определяется добротностью резонаторов.

Также следует учесть, что с ростом рабочей частоты размеры шлейфов значительно уменьшаются. Например, на частотах свыше 10 ГГц размеры четвертьволновых шлейфов становятся столь малы, что их использование для согласования становится крайне затруднительным.

Кольцевой эллиптический резонатор

В качестве альтернативы можно рассмотреть возможность применения кольцевых эллиптических резонаторов (КЭР) для создания различных микроволновых устройств. КЭР представляет собой полосковую линию, изогнутую в кольцо и соединенную с основной линией передачи (рисунок 2).

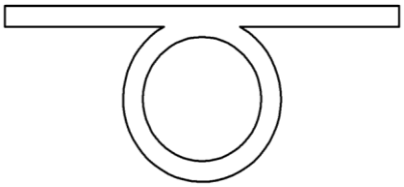


Рис. 2. Топология простого КЭР

Fig. 2. The Topology of Simple Circular Elliptical Resonator (CER)

Выражения (3) и (4) отображают зависимость резонансной частоты и геометрии кольца. Взаимосвязь резонансной частоты с эквивалентными индуктивностью и емкостью кольца устанавливает в соответствии с (5).

$$f_{\text{рез1}} = \frac{c}{\lambda\sqrt{\varepsilon}}, \quad (3)$$

$$l_{\text{ср}} = \left(1 + \frac{n}{2}\right)\lambda, \quad (4)$$

$$f_{\text{рез2}} = \frac{1}{2\pi\sqrt{LC}}. \quad (5)$$

Модель КЭР и результаты ее расчета представлены на рисунках 3 и 4.

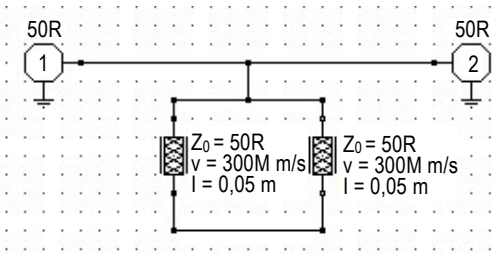


Рис. 3. Модель КЭР

Fig. 3. The Model of Simple CER

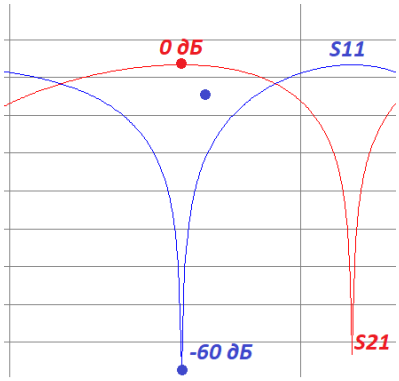


Рис. 4. Общий вид *S*-параметров КЭР

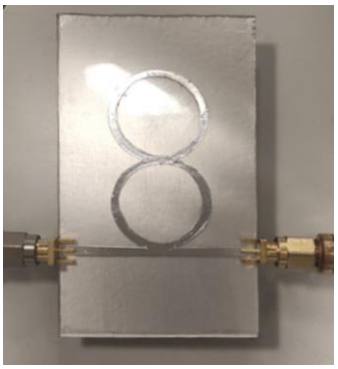
Fig. 4. The Image of S-Parameters Characteristics

Резонансная частота КЭР непосредственно связана с его геометрическими параметрами. Она определяется количеством длин полуволн, укладывающихся в кольцо (3, 4), с учетом диэлектрической проницаемости материала. Кроме того, КЭР обладает и другой резонансной частотой, связанной с его погонными параметрами. Если представить КЭР как индуктивность и емкость линии передачи, то резонансную частоту можно вычислить по формуле (5).

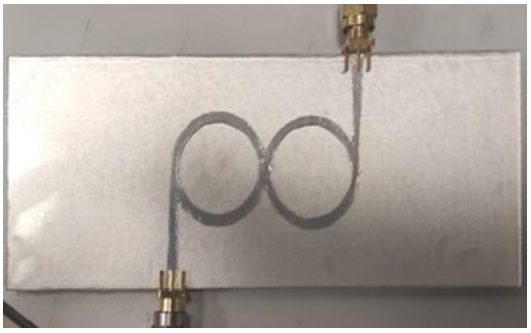
Исследованию резонансных свойств КЭР в зависимости от геометрии топологии посвящена статья [1], в которой была доказана работоспособность КЭР. Ключевым достижением стало совмещение двух ранее рассмотренных резонансных частот, что позволило достичь ослабления на резонансной частоте более 30 дБ. Стоит отметить, что такое значительное ослабление радиосигнала стало возможным только благодаря использованию кольца.

Двойной кольцевой эллиптический резонатор

В рамках эксперимента конструкция двойного КЭР была успешно интегрирована в макет платы СВЧдиапазона. Были изготовлены макеты КЭР с расчетной частотой 3 ГГц (рисунок 5) и проведены измерения на векторном анализаторе цепей в частотном диапазоне от 1 до 15 ГГц.



a)



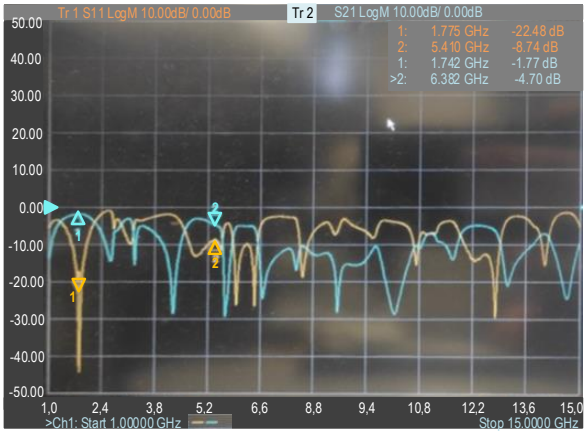
b)

Рис. 5. Макеты двойного КЭР: a) типа «восьмерка»; b) типа «спираль»

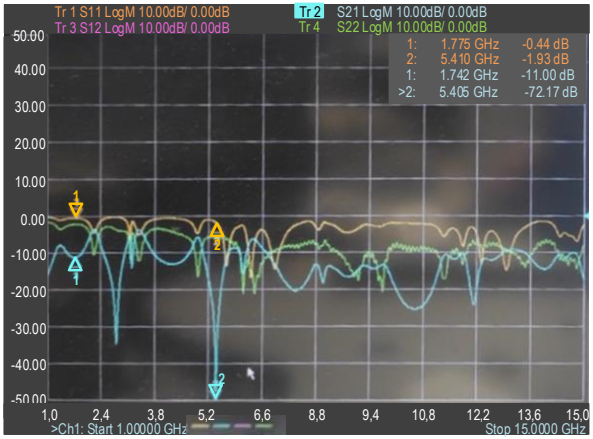
Fig. 5. Double CER Layouts in the Form of "Eight" (a) and "Spiral" (b)

В ходе эксперимента были исследованы различные способы включения резонаторов в линию тракта, а также рассмотрено, как эти способы влияют на итоговые характеристики резонаторов.

Результаты исследования первого макета (рисунок 6a) показали, что конструкция с подключением типа «восьмерка» обеспечивает наилучшее согласование с точки зрения потерь на частоте около 1,5 ГГц. Но наибольший интерес представляют данные эксперимента с топологией макета «спираль» (рисунок 6b). На частоте 2,7 ГГц, которая была близка к расчетной, был замечен резонанс с ослаблением более 30 дБ. Но самое удивительное – это вторая резонансная частота характеристики, которая была вдвое выше первой и составляла примерно 5,4 ГГц, с ослаблением около 72 дБ.



a)



b)

Рис. 6. Результаты эксперимента двойного КЭР типа «восьмерка» (a) и типа «спираль» (b)

Fig. 6. Measurement Results CER in the Form of "Eight" (a) and "Spiral" (b)

За годы научной работы в СПбГУТ исследовались КЭР различных конструкций: как планарные, так и объемные структуры кольцевых резонаторов, а также всевозможные типы питающих линий (микрополосковые, спиральные, коаксиальные, объемный волновод) [2, 3]. Тема КЭР стала предметом интереса ученых из различных областей науки [4–6]. После проведения многочисленных исследований было вновь подтверждено, что свойства устройства в целом во многом зависят от способа питания кольца.

Преселектор частоты на основе КЭР

Из сказанного ранее следует, что один КЭР можно легко использовать для создания преселективных фильтров, которые будут работать на частоте тракта. Также существует возможность объединить несколько КЭР с разными резонансными частотами в одно устройство, которое будет функционировать на близких частотах. Ожидается, что с помощью одного канала передачи можно будет обрабатывать сигналы на разных частотах, и эта структура будет напоминать аналоговый демультиплексор.

Усилитель на КЭР

Несмотря на свою простоту, КЭР обладает значительным потенциалом. Помимо своих селективных свойств, он может применяться для усиления и сложения мощности активных двухполюсников.

В работе [7] были исследованы различные способы включения активных двухполюсников в структуру кольцевого резонатора. В одном из экспериментов использовался обычный туннельный диод, соединенный с кольцом. В итоге было достигнуто усиление сигнала на заданной частоте до 14 дБ при конкретной геометрии КЭР. Этот результат может найти применение в устройствах с подавлением побочных гармоник.

Аналогичная конструкция КЭР также рассматривалась как способ суммирования мощности сигналов, исходящих от нескольких активных двухполюсников [8].

Генератор на основе КЭР

Классическая структура микроволнового генератора представляет собой источник электромагнитных колебаний, окруженный резонатором. В работах [9–11] были изучены генераторы на основе КЭР с интегрированным двухполюсником. В ходе эксперимента была достигнута впечатляющая согласованность результатов расчетов и эксперимента (рисунки 7–9).

Проблема коммутации КЭР и основного тракта

В последнее время ученые активно изучают распределение электромагнитного поля в КЭР. Как уже неоднократно отмечалось, свойства КЭР во многом зависят от способа питания резонатора. При обычной гальванической связи волна входит в кольцо и распространяется в двух направлениях, что приводит к образованию стоячей волны. Теоретически, можно создать условия для распространения волны только в одну сторону, что значительно повысит добротность резонатора. Создание такого

резонатора позволит создать бегущую волну, потенциал использования которой практически безграничен.

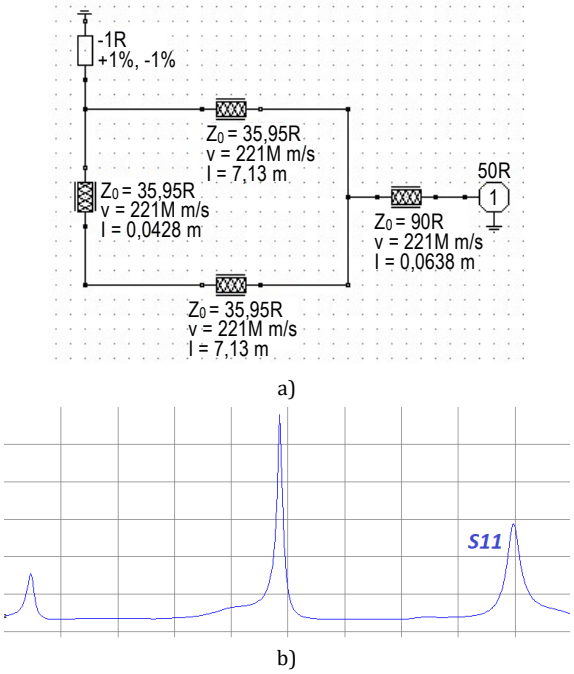


Рис. 7. Принципиальная схема (a) и эмуляция работы (b) генератора на КЭР в программе RFSimm

Fig. 7. RFSimm Schematic Diagram (a) and RFSimm Simulation (b) of the Generator Based on the CER

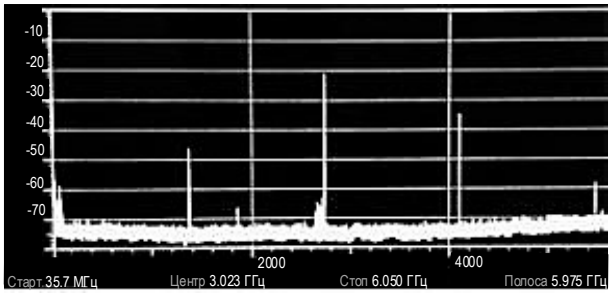


Рис. 8. Результаты эксперимента по генерации частоты на КЭР

Fig. 8. Experimental Results of the Generator Based on CER

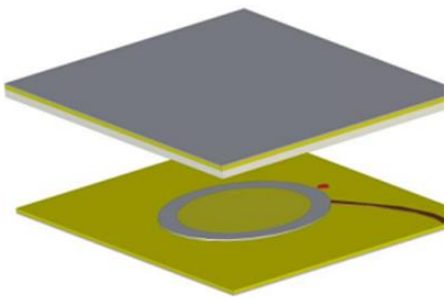


Рис. 9. Эпюр макета генератора на КЭР

Fig. 9. Plot of Generator Based on CER

В [7] были рассмотрены несколько способов ограничения направлений распространения электромагнитной волны в КЭР: использование ферритовых вентилей и направленных ответвителей.

В ходе исследования были разработаны несколько моделей КЭР, в которые были интегрированы микроволновые вентили (рисунок 10). В результате эксперимента было достигнуто невзаимное распространение волны в тракте. Потери на рабочей частоте от порта 1 к порту 2 составили 15 дБ, а от порта 2 к порту 1 – около 21 дБ. Этот результат свидетельствует о том, что удалось частично ограничить направление распространения волны.

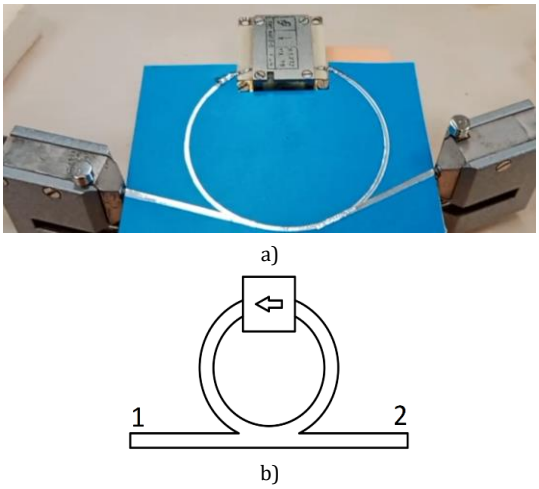


Рис. 10. Макет КЭР (a) и его структура (b) с интегрированным вентилем

Fig. 10. The CER Layout (a) and its Structure (b) with an Integrated RF-Valve

Также был рассмотрен метод возбуждения волны в резонаторе с помощью направленных ответвителей на связанных линиях с различной направленностью (рисунок 11). Было исследовано несколько макетов, в которых кольцо возбуждалось через обычные ответвители и направленные ответвители на нерегулярной линии. Эксперимент показал, что масштабный макет с направленными осветителями на обычных связанных линиях имеет большее ослабление (около –25 дБ), чем макет с ответвителями на нерегулярных линиях (около 12 дБ).

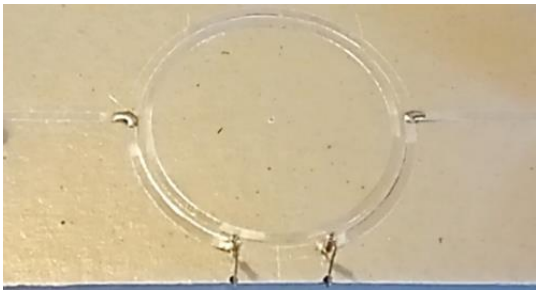


Рис. 11. Макет КЭР с возбуждение волны через направленные ответвители

Fig. 11. A CER Layout with Wave Excitation through Directional Couplers

Заключение

Для создания КЭР можно использовать различные типы линий: микрополосковые, коаксиальные, копланарные, щелевые и другие. Эти резонаторы нашли широкое применение в самых различных устройствах: от частотной селекции до усилителей, генераторов, устройств распределения мощности и даже в создании не взаимных систем. Большинство макетов КЭР изготовлено с использованием аппликационного метода на органических диэлектриках, их работоспособность была доказана экспериментально. Применение промышленных материалов и технологий изготовления, безусловно, позволит улучшить характеристики устройств на основе КЭР. Применение резонаторов подобного типа расширяет возможности разработчиков и создает дополнительную элементную базу интегральной микроволновой электроники. Добротность, технологичность, конформность и удобство подстройки делают КЭРы незаменимыми в современной микроэлектронике.

Список источников

1. Сазоненко Н.Ю., Седышев Э.Ю. Устройства частотной селекции на основе кольцевых эллиптических резонаторов на микрополосковой линии // Электроника и микроэлектроника СВЧ. 2019. Т. 1. С. 409–411. EDN:NVXLXX

2. Кондрашова М. А., Сазоненко Н. Ю., Селиверстов Л. А., Улитина А. С., Седышев Э. Ю. Частотно-селективные устройства на кольцевых эллиптических резонаторах // Проектирование и технология электронных средств.2019. № 2. С. 13-20. EDN:XCNOKU

3. Леонтьев А.С., Седышев Э.Ю. Синтез устройства частотной селекции на кольцевых эллиптических резонаторах в объёмном интегральном исполнении // Электроника и микроэлектроника СВЧ. 2022. Т. 1. С. 382–386. EDN:SYGCBG

4. Коркина А.Р. Микроволновый датчик для анализа примесей в оливковом масле // Международная молодёжная научная конференция, посвященная 60-летию со дня осуществления Первого полета человека в космическое пространство и 90-летию Казанского национального исследовательского технического университета им. А.Н. Туполева-КАИ «XXV Туполевские чтения (школа молодых ученых)» (Казань, Российская Федерация, 10–11 ноября 2021 г.). Т. VI. Казань: ИП Сагиева А.Р., 2021. С. 166–171. EDN:GRUMXI

5. Lobekin V., Tatarenko A., Belyshev A., Bichurin M. Resonator for micro-wave magnetoelectric effect // Proceedings of the 29th International Crimean Conference “Microwave & Telecommunication Technology” (CriMiCo’2019, Sevastopol, Russian Federation, 8–14 September 2019). 2019. Vol. 30. DOI:10.1051/itmconf/20193007012

6. Коркина А.Р., Насыбуллина А.Р., Фархутдинов Р.В. Объемные кольцевые резонаторы в копланарном исполнении в качестве СВЧ датчиков для определения диэлектрической проницаемости жидкостей // XXIV Международная научно-техническая конференция и материалы XX Международной научно-технической конференции «Проблемы техники и технологии телекоммуникаций. Оптические технологии в телекоммуникациях» (Уфа, Российская Федерация, 23–25 ноября 2022 г.). Уфа: Федеральное государственное бюджетное образовательное учреждение высшего образования "Уфимский университет науки и технологий", 2023. Т. 1. С. 345–347. EDN:COSIKS

7. Иванищева Е.Ф., Леонтьев А.С., Седышев Э.Ю., Федоров С.И. Особенности распространения волны в кольцевом эллиптическом резонаторе // Электроника и микроэлектроника СВЧ. 2024. № 1. С. 521–525. EDN:EZOWXQ

8. Бочаров Е.И., Подольская М.О., Седышев Э.Ю. Усилитель на активном двухполюснике, интегрированный в кольцевой эллиптический резонатор // IX Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (АПИНО-2020, Санкт-Петербург, Российская Федерация, 26–27 февраля 2020 г.). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2020. Т. 3. С. 408–412. EDN:TSNVZG

9. Седышев Э.Ю., Шомин А.Ю., Исследование возможности одновременного использования нескольких активных двухполюсников при создании СВЧ генераторов // IX Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (АПИНО-2020, Санкт-Петербург, Российская Федерация, 26–27 февраля 2020 г.). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2020. Т. 3. С. 514–519. EDN:NNESFL

10. Шомин А.Ю., Седышев Э.Ю. Генератор СВЧ в интегральном исполнении на кольцевом резонаторе // Региональная научно-методическая конференция магистрантов и их руководителей (Санкт-Петербург, Российская Федерация, 01–03 декабря 2020 г.). Подготовка профессиональных кадров в магистратуре для цифровой экономики (ПКМ-2020). Сборник лучших докладов конференции. СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2021. С. 339–343. EDN:HEOMTO

11. Каткова Т.О., Седышев Э.Ю., Генератор СВЧ на кольцевом эллиптическом резонаторе в объемном интегральном исполнении // Электроника и микроэлектроника СВЧ. 2021. Т. 1. С. 430–433. EDN:SWJDTM

References

1. Sazonenko N.Yu., Sedyshev E.Yu. Frequency Selection Devices Based on Circular Elliptical Resonators on a Microstrip Line. *Elektronika i mikroelektronika SVCH*. 2019;1:409–411. (in Russ.) EDN:NVXLXX

2. Kondrashova M.A., Sazonenko N.Y., Seliverstov L.A., Ulitina A.S., Sedyshev E.Y. Frequency-selective devices on circular elliptical resonators. *Design and Technology Of Electronic Devices*.2019;2:13–20. (in Russ.) EDN:CSGJRK

3. Leontyev A.S., Sedyshev E.Yu. Synthesis of a frequency selection device based on ring elliptic resonators in a three-dimensional integral design. *Elektronika i mikroelektronika SVCH*. 2022;1:382–386. (in Russ.) EDN:SYGCBG

4. Korkina A. Microwave sensor for analysis of impurities in olive oil. *Proceedings of the International Youth Scientific Conference dedicated to the 60th Anniversary of the First Manned Flight into Space and the 90th Anniversary of the Kazan National Research Technical University named after A.N. Tupolev-KAI, 10–11 November 2021, Kazan, Russian Federation. XXV Tupolev Readings. School of Young Scientists*. Kazan: Sagieva A.R. Publ.; 2021. p.166–171. (in Russ.) EDN:GRUMXI

5. Lobekin V., Tatarenko A., Belyshev A., Bichurin M. Resonator for micro-wave magnetoelectric effect. *Proceedings of the 29th International Crimean Conference "Microwave & Telecommunication Technology", CriMiCo'2019, 8–14 September 2019, Sevastopol, Russian Federation, vol.30*. 2019. DOI:10.1051/itmconf/20193007012

6. Korkina A.R., Nasybullin A.R., Farkhutdinov R.V. Volume Ring Resonators in a Coplanar Design as Microwave Sensors for Determination of the Dielectric Permittivity. *Proceedings of the XXIV International Scientific and Technical Conference and Proceedings of the XXth International Scientific and Technical Conference "Problems of Telecommunication Engineering and Technology. Optical Technologies in Telecommunications", 23–25 November 2022, Ufa, Russian Federation*. Ufa: Ufa University of Science and Technology Publ.; 2023. p.345–347. (in Russ.) EDN:COSIKS

7. Ivanishcheva E.F., Leontyev A.S., Sedyshev E.Y., Fedorov S.I. Features of wave propagation in an elliptical ring resonator. *Elektronika i mikroelektronika SVCH.* 2024;1:521–525. (in Russ.) EDN:EZOWXQ

8. Bocharov E., Podolskaya M., Sedyshev E. Amplifier on an Active Two-Pole Integrated into a Ring Elliptical Resonator. *Proceedings of the IXth International Scientific, Technical and Scientific-Methodological Conference "Actual Problems of Infotelecommunications in Science and Education", APINO-2020, 26–27 February 2020, St. Petersburg, Russian Federation, vol.3*. St. Petersburg: The Bonch-Bruevich Saint Petersburg State University of Telecommunications Publ.; 2020. p.408–412. (in Russ.) EDN:TSNVZG

9. Sedyshev E., Shomin A. Research of the Possibility of Simultaneous Use of Several Active One-Port Devices in Creation of Generators. *Proceedings of the IXth International Scientific, Technical and Scientific-Methodological Conference "Actual Problems of Infotelecommunications in Science and Education", APINO-2020, 26–27 February 2020, St. Petersburg, Russian Federation, vol.3*. St. Petersburg: The Bonch-Bruevich Saint Petersburg State University of Telecommunications Publ.; 2020. p.514–519. (in Russ.) EDN:NNESFL

10. Shomin A., Sedyshev E. The Microwave Generator in Integral Design on a Ring Resonator. *Regional Scientific and Methodological Conference of Undergraduates and Their Supervisors, 01–03 December 2020, St. Petersburg, Russian Federation. Training of Professional Personnel in the Magistracy for the Digital Economy. Collection of the Best Conference Reports*. St. Petersburg: The Bonch-Bruevich Saint Petersburg State University of Telecommunications Publ.; 2021. p.339–343. (in Russ.) EDN:HEOMTO

11. Katkova T.O., Sedyshev E.Yu. Microwave Generator on an Elliptical Ring Resonator in a Three-Dimensional Integral Design. *Elektronika i mikroelektronika SVCH.* 2021;1:430–433. (in Russ.) EDN:SWJDTM

Статья поступила в редакцию 25.11.2024; одобрена после рецензирования 19.12.2024; принята к публикации 23.12.2024.
The article was submitted 25.11.2024; approved after reviewing 19.12.2024; accepted for publication 23.12.2024.

Информация об авторах:

ЛЕОНТЬЕВ Александр Сергеевич — аспирант, ассистент кафедры Электроники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0009-0007-5778-5274

СЕДЫШЕВ Эрнест Юрьевич — кандидат технических наук, доцент кафедры Электроники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0009-0008-0002-3049

Авторы сообщают об отсутствии конфликтов интересов.
The authors declare no conflicts of interests.

Научная статья
УДК 621.396.67
https://doi.org/10.31854/1813-324X-2024-10-6-34-44

Методы пространственной обработки спутниковых навигационных сигналов в частотной области

Владимир Игоревич Царик, wladimirzarik@mail.ru

ООО «Эйртэго»,
Санкт-Петербург, 197375, Российская Федерация.

Аннотация

Актуальность. *Весьма низкая мощность полезных информационных сигналов глобальных спутниковых навигационных систем вблизи поверхности Земли вместе с происходящим в последние годы заметным увеличением количества доступных и эффективных портативных средств постановки заградительных широкополосных энергетических помех делают задачу повышения помехоустойчивости радионавигационных спутниковых устройств особенно актуальной как с практической, так и с исследовательской точек зрения. В этой связи* ***целью*** *данного исследования явилось повышение помехоустойчивости глобальных спутниковых навигационных систем посредством обработки входных сигналов соответствующей принимающей аппаратуры специальными пространственными фильтрами. Для достижения цели работы была решена научная задача по исследованию увеличения помехоустойчивости радионавигационной аппаратуры с использованием в ней пространственной обработки входных сигналов в частотной области.*

Используемые методы. *В ходе исследования были рассмотрены различные алгоритмы пространственной обработки сигналов, среди которых были как функционирующие в условиях отсутствия какой-либо информации о внешней относительно принимающей радионавигационной системы помеховой обстановке, так и задействующие сведения о количестве и относительном расположении источников помех. Дополнительно были исследованы различные методы нахождения числа источников помех и угловых направлений на них, а также современные алгоритмы оптимизации целевых функций, используемых для определения местоположения источников сигналов.*

Научная новизна *работы заключается в применении при решении поставленной задачи новых алгоритмов, реализующих отдельные этапы сигнальной обработки и обеспечивающих получение алгоритмами фильтрации информации, необходимой для их работы, а также в комбинировании известных методов с новыми подходами к их воплощению.*

Результаты. *В ходе решения научной задачи было проведено сравнение характеристик качества работы всех рассмотренных алгоритмов, выполненное с применением метода компьютерного моделирования, при котором использовались записи реальных спутниковых навигационных сигналов с добавлением разного количества источников некоррелированных энергетических помех. В результате моделирования были получены значения показателей качества работы всех исследуемых алгоритмов и проведен их сравнительный анализ, по итогам которого выделены методы с наилучшими характеристиками.*

Значимость *результатов работы состоит в возможности использовать рассмотренные алгоритмы при разработке реальных устройств помехозащищенной спутниковой навигации.*

Ключевые слова: *спутниковые навигационные системы, помехоустойчивость, энергетическая помеха, пространственная обработка, частотная область, MATLAB*

Ссылка для цитирования: Царик В.И. Методы пространственной обработки спутниковых навигационных сигналов в частотной области // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 34–44. DOI:10.31854/1813-324X-2024-10-6-34-44. EDN:VINYXC

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-34-44

Space-Frequency Processing Methods for Satellite Navigation Signals

Vladimir I. Tsarik, wladimirzarik@mail.ru

Airtago LLC,
St. Petersburg, 197375, Russian Federation

Annotation

Relevance. *Quite low power of the global satellite navigation systems' useful informational signals near the Earth surface along with an ongoing noticeable increase of the number of easily available and efficient portable means of blocking wideband energetic interference radiation make the problem of radionavigational satellite devices antijamming capabilities improvement especially relevant both from practical and scientific points of view. Therefore,* ***the goal of this research*** *was to increase the antijamming capabilities of the global satellite navigation systems via processing of the corresponding receiving apparatus' input signals by special spatial filters. To achieve the work goal the scientific task of researching on the antijamming capability improvement in radionavigational devices by means of space-frequency signal processing was solved.*

The methods used. *During the research, different spatial signal processing algorithms were considered, among them both the ones functioning without any information about interference situation, external with respect to the receiving radionavigational system, and the ones using the knowledge about the number and relative disposition of the jamming sources. Additionally different methods of interference sources number and angular directions finding were studied, as well as modern cost function optimization algorithms which are used for signal sources' location determination.*

Scientific novelty *of this work consists of usage of new algorithms that implement separate signal processing stages and that provide necessary information to the filtering algorithms during the problem solution, as well as of combining known methods with new approaches to their design.*

The results. *During the scientific task solution, the performance quality metrics comparison was carried out for all the considered algorithms via the computer modeling method that employed recordings of real satellite navigational signals with addition of varying number of uncorrelated energetic interferences sources. As a result of modeling, the performance quality measure values were obtained for all the investigated algorithms and the comparative analysis thereof was conducted, at the end whereof the methods with the best characteristics were picked out.*

The significance *of the work results consists of possibility of using the considered algorithms in real antijamming satellite navigation devices design.*

Keywords: *satellite navigation systems, antijamming capability, energetic interference, spatial processing, frequency domain, MATLAB*

For citation: Tsarik V.I. Space-Frequency Processing Methods for Satellite Navigation Signals. *Proceedings of Telecommunication Universities.* 2024;10(6):34–44. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-34-44. EDN:VINYXC

Введение

Сигналы глобальных навигационных спутниковых систем (ГНСС), широко применяющихся в многих критически важных областях человеческой деятельности, обладают весьма малой мощностью вблизи поверхности Земли [1]. Вследствие этого радионавигационные спутниковые системы чрезвычайно уязвимы к воздействию различного рода помех. К ним относятся как помехи, возникающие в технических системах естественным образом, так и внешние воздействующие сигналы, обычно излучаемые преднамеренно с целью нарушения правильности работы радионавигационной системы. Одним из наиболее эффективных видов такого воздействия является широкополосная заградительная энергетическая помеха, которая в максимальной степени заполняет рабочий диапазон частот шумоподобными сигналами с вы-

соким уровнем, маскируя собой полезные информационные сигналы [2]. В последние годы особенно сильно выросло число доступных средств излучения искусственных заградительных помех, а также опасных происшествий, связанных с их использованием. Все это делает достаточно актуальной задачу повышения помехоустойчивости приемников сигналов ГНСС.

Одним из известных и действенных способов решения данной задачи является пространственная обработка сигналов ГНСС в частотной области (SFAP, *аббр. от англ*. Space-Frequency Adaptive Processing). Данный вид обработки реализуется внутри принимающей аппаратуры спутниковых сигналов, использующей многоэлементные антенные решетки (АР) и задействует сигналы, уже прошедшие через радиочастотную часть приемного тракта системы и ее аналого-цифровые преобразователи (АЦП). По сравнению с обычной пространственной обработкой сигналов (ПОС), фильтрация в частотной области обладает рядом преимуществ при незначительных отличиях в вычислительной сложности реализации [3].

В данной работе рассматриваются несколько алгоритмов ПОС в частотной области. Часть из них способна функционировать в условиях отсутствия каких-либо априорных сведений о сигнально-помеховой обстановке, в которой находится принимающая спутниковые сигналы радиотехническая система. Другие используют такую информацию о присутствующих вблизи системы помехах, как их число и относительные угловые направления на них. Помимо этого, в работе также рассматриваются различные алгоритмы, сопутствующие обработке данных с использованием информации об окружающей среде. К ним относятся методы оценки числа действующих источников помех, определения направлений на них, а также алгоритмы оптимизации функций, описывающих пространственную конфигурацию источников сигналов, окружающих приемную радионавигационную систему. В ходе работы проводится описание рассматриваемых методов, а также компьютерное моделирование с их применением и использованием записей реальных сигналов от ГНСС и помех, по итогам которого выполняется сравнительный анализ различных показателей качества работы всех рассмотренных алгоритмов. Данное исследование обобщает и углубляет результаты, описанные автором в более ранних работах по этой теме [4, 5].

Повышение помехоустойчивости радионавигационных систем

Задача повышения помехоустойчивости радионавигационного приемника сигналов ГНСС рассматривается в данной работе в следующей постановке. АР, состоящая из N антенных элементов (АЭ), расположена в плоскости Oxy. Над АР в верхнем полупространстве находятся источники полезного сигнала – навигационные спутники – и N_I попарно некоррелированных источников широкополосных энергетических помех. Упрощенное схематическое изображение данной ситуации приведено на рисунке 1. Поступающий на АР сигнал представляет собой аддитивную смесь полезного спутникового сигнала, помех и белого гауссовского шума. При этом уровень полезного сигнала, как отмечалось выше, не превосходит уровень шума, а суммарный уровень помех, наоборот, значительно его превосходит. После прохождения приемных трактов радионавигационного устройства и АЦП поступивший на АР сигнал принимает вид дискретных комплексных отсчетов. При рассмотрении записи данного сигнала, имеющей длину L, удобно представить его в виде комплексной матрицы x, соответствующей размерам N и L. Требуется построить процедуру обработки входного сигнала x, преобразующей его в выходной сигнал y, в котором в максимальной степени подавлены помехи и достаточно высокое качество для успешного решения навигационной задачи с его использованием.

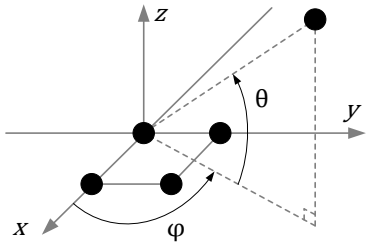


Рис. 1. Схема взаимного расположения АР и источника сигнала

Fig. 1. Antenna Array and a Signal Source Mutual Placement Scheme

Уровень качества обработанного сигнала и степень подавления в нем помех можно выразить с помощью различных энергетических характеристик помехоустойчивости. Для навигационных сигналов одним из важнейших критериев качества является итоговая точность позиционирования с их использованием. Можно показать, что дисперсия ошибок позиционирования по конкретному сигналу зависит от его отношения сигнал/шум (ОСШ), которое равно отношению мощности полезного сигнала к мощности шума, обычно выражаемому в децибелах. Данная зависимость обратно пропорциональна, то есть повышение ОСШ влечет за собой уменьшение ошибок позиционирования [6]. В этой связи данный энергетический показатель помехоустойчивости сигнала можно использовать в качестве характеристики его качества после выполнения процедуры обработки.

В качестве показателя степени подавления помех в результате фильтрации можно применять

коэффициент подавления (КП) помех, равный отношению мощностей сигнала до и после обработки [7]. Данное значение характеризует долю присутствующих в сигнале помех, успешно подавленных в результате обработки. Повышение этого показателя помехоустойчивости очевидным образом увеличивает ОСШ результирующего сигнала и, вследствие этого, также способствует улучшению качества итогового позиционирования.

Пространственная обработка сигналов в частотной области

В работе в качестве процедуры обработки спутниковых сигналов, повышающей помехоустойчивость соответствующих навигационных приемников, была рассмотрена пространственная обработка в частотной области. В основе данного способа фильтрации лежат методы ПОС, которые преобразуют входные сигналы АР с использованием информации о ее пространственной конфигурации. Такая обработка более эффективна, чем применение обычных фильтров с конечной импульсной характеристикой, но в то же время она обладает определенными недостатками. Наиболее важным из них является ограничение количества потенциально подавляемых источников помех числом, на единицу меньшим количества используемых АЭ. Применение методов ПОС к сигналам в частотной области позволяет не только снять данное ограничение, но и существенно увеличить общее качество обработки при минимальном увеличении вычислительных затрат [3].

Общая схема ПОС в частотной области приведена на рисунке 2, где $x_1, \dots, x_N$ – строки матрицы x, обозначающие N входных сигналов АР; ДПФ – дискретное преобразование Фурье, задающееся формулой [8]:

$$X(k) = \sum_{j=1}^{N} x(j) e^{-\frac{2\pi i}{N}(j-1)(k-1)}, k = 1, \dots, N,$$

где i – мнимая единица; X и Y – Фурье-образы, соответственно, сигналов x и y.

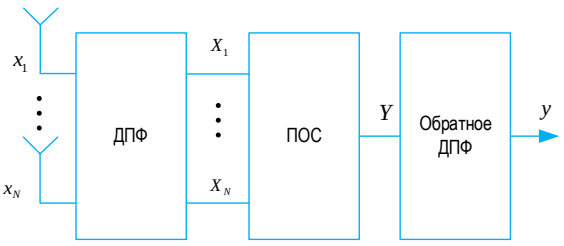


Рис. 2. Блок-схема пространственной обработки в частотной области

Fig. 2. Flowchart of the Space-Frequency Adaptive Processing

Наиболее эффективной данная схема обработки будет являться в случае, когда длина записи N равна степени двойки. Далее без умаления общности предполагается, что такое равенство имеет место. В данном случае для вычисления ДПФ можно пользоваться специальными быстрыми алгоритмами [9]. Дополнительно увеличить эффективность и адаптивность обработки по представленной схеме и облегчить ее потенциальную реализацию в существующих вычислительных устройствах можно путем выполнения ДПФ не над всей записью сигнала, а над ее последовательными секциями по M отсчетов каждая, где число M является степенью двойки, которая делит N. Дальнейшую ПОС также удобно выполнять не над всей секцией длины M, а над ее последовательными сегментами длины T, где число T, в свою очередь, является степенью двойки, делящей M [10].

В каждом сегменте частотные отсчеты Y выходного сигнала вычисляются посредством умножения частотных отсчетов X входного сигнала на вектор весовых коэффициентов (с) w пространственного фильтра:

$$Y = w^H X,$$

где верхний индекс H обозначает эрмитово сопряжение.

Затем по частотным отсчетам Y выходного сигнала посредством обратного ДПФ вычисляются соответствующие временны́е отсчеты y, и далее вся процедура обработки повторяется для каждого сегмента каждой секции отсчетов входного сигнала.

В следующем подразделе будут описаны алгоритмы ПОС, использовавшиеся в данной работе для фильтрации в частотной области.

Методы пространственной обработки сигналов

Алгоритмы ПОС, рассматриваемые ниже, можно разделить на две группы: методы, не задействующие информацию об окружающей принимающую систему сигнально-помеховой обстановке, и методы, нуждающиеся в таких данных, в частности, в количестве источников действующих вблизи системы помех и направлениях на них. Алгоритмы первого рода более просты в реализации, но обладают меньшей способностью к адаптации обработки к постоянно изменяющимся внешним условиям по сравнению с методами второго класса.

В качестве алгоритма первой группы в данной работе был использован метод обработки, известный как бимформер Кейпона (*от англ.* Capon) или MPDR-бимформер (*аббр. от англ.* Minimum Power Distortionless Response – неискаженный отклик наименьшей мощности).

ВВК данного фильтра задается выражением:

$$w_{\mathrm{MPDR}}(\theta, \varphi) = \frac{R^{-1} a(\theta, \varphi)}{a^H(\theta, \varphi) R^{-1} a(\theta, \varphi)}, \quad (1)$$

где R – корреляционная матрица (КМ) входного сигнала АР; $a(\theta,\varphi)$ – управляющий вектор АР в направлении, заданном азимутом θ и углом места φ:

$$a(\theta,\varphi)=\exp\left\{i\frac{2\pi}{\lambda}uv(\theta,\varphi)\right\}, \quad (2)$$

λ – длина волны приходящего на АР полезного сигнала; $u\in\mathbb{R}^{N\times 3}$ – матрица из декартовых координат АЭ; $v(\theta,\varphi)$ – вектор-столбец единичной нормы, выражающий направление, определенное углами θ и φ, в декартовых координатах [11]:

$$v(\theta,\varphi)=\begin{pmatrix}\cos\theta\cos\varphi\\ \cos\theta\sin\varphi\\ \sin\theta\end{pmatrix}.$$

Необходимые для обработки значения углов θ_* и φ_*, которые в данном случае считаются априори неизвестными, можно определить по результату обработки сигнальной выборки x' длины $L_S<L$ посредством максимизации значения КП, то есть из выражения:

$$(\theta_*,\varphi_*)=\arg\max_{\theta,\varphi,i}\left\{\frac{\mathbb{D}x_i'}{\mathbb{D}\{w(\theta,\varphi)x'\}}\right\}, \quad (3)$$

где $\mathbb{D}\{\cdot\}$ – оператор вычисления выборочной дисперсии.

Описанная процедура позволяет проводить ПОС в условиях отсутствия информации о количестве присутствующих вблизи радиотехнической системы источников помехи и их расположении в пространстве, а также при отсутствии необходимости или возможности получения такой информации.

Одним из рассматриваемых в данной работе алгоритмов ПОС, использующих априорную информацию об источниках помех, является так называемый LCMP-бимформер (*аббр. от англ.* Linear Constrained Minimum Power – наименьшая мощность с линейными ограничениями). Пусть уже известны количество источников помех N_I и соответствующие направлениям на них углы – азимуты $\theta_1,\dots,\theta_{N_I}$ и углы места $\varphi_1,\dots,\varphi_{N_I}$. Тогда по ним с помощью выражения (2) можно построить N_I управляющих векторов:

$$a_j=a\left(\theta_j,\varphi_j\right), j=1,\dots,N_I,$$

с использованием которых затем определить матрицу:

$$S(\theta,\varphi)=\left(a(\theta,\varphi),a_1,\dots,a_{N_I}\right),$$

где $a(\theta,\varphi)$ – управляющий вектор с переменными значениями углов θ и φ, также определяемый формулой (2).

С помощью вектора-столбца $C=(1,0,\dots,0)^T$, в котором количество нулей равно N_I, можно определить ВВК LCMP-фильтра как:

$$w_{\text{LCMP}}=R^{-1}S(S^HR^{-1}S)^{-1}C. \quad (4)$$

Сформулированное при постановке задачи повышения помехоустойчивости предположение о том, что шум во входном сигнале является белым, позволяет упростить выражение (4) до ВВК так называемого LCMV-бимформера (*аббр. от англ.* Linear Constrained Minimum Variance – наименьшая дисперсия с линейными ограничениями), который имеет следующий вид [11]:

$$w_{\text{LCMV}}=S(S^HS)^{-1}C.$$

Для обоих формирователей лучей с линейными ограничениями оптимальные значения углов θ и φ также определяются посредством максимизации КП в соответствии с формулой (3).

Алгоритмы определения количества источников помех

В основе всех использовавшихся в рамках данной работы методов определения количества источников помех лежит следующая идея. Предполагается, что множество всех сигналов, допустимых в приведенной выше постановке задачи, является конечномерным линейным пространством, которое можно разложить в прямую сумму двух собственных подпространств, содержащих соответственно полезные сигналы и помехи.

В этой связи множество $\Lambda=\{\lambda_1,\dots,\lambda_N\}$ собственных чисел КМ входного сигнала также распадается на два подмножества, соотносящиеся указанным подпространствам, причем N_I собственных значений из подмножества, соответствующего помеховому подпространству, оказываются больше, чем остальные $N_S=N-N_I$ собственных чисел.

Этот факт можно описать следующей цепочкой соотношений:

$$\lambda_{N_I}^I\geq\dots\geq\lambda_1^I>\lambda_{N_S}^S\approx\dots\approx\lambda_1^S.$$

Вследствие этого задача определения количества источников помех сводится к подсчету количества достаточно больших собственных чисел КМ входного сигнала [12].

Одним из наиболее часто используемых и проработанных подходов к решению данной задачи является использование методов статистического последовательного анализа.

Значительным результатом данной области знаний, полученным при решении этой задачи, является тот факт, что точка минимума функции:

$$F(d)=K(N-d)\ln\left\{\frac{\frac{1}{N-d}\sum_{j=d+1}^{N}\lambda_j}{\left(\prod_{j=d+1}^{N}\lambda_j\right)^{\frac{1}{N-d}}}\right\},$$

где K – количество отсчетов сигнала, используемых для вычислений; $d=0,\dots,N-1$, может служить оценкой числа достаточно больших значе-

ний в наборе собственных чисел Λ. Этот факт лежит в основе двух наиболее часто используемых для решения данной задачи методов.

В соответствии с одним из них – информационным критерием Акаике (AIC, *аббр. от англ.* Akaike Information Criterion) – в качестве оценки количества источников помех используется значение:

$$N_{\text{AIC}} = \arg\min_{d}\{F(d) + d(2N - d)\}.$$

При применении другого алгоритма – метода наименьшей описательной длины (MDL, *аббр. от англ.* Minimum Description Length) – для оценки этого числа используется значение:

$$N_{\text{MDL}} = \arg\min_{d}\left\{F(d) + \frac{1}{2}(d(2N - d) + 1)\ln K\right\}.$$

Известно, что при неограниченном увеличении K оценка метода MDL становится состоятельной, в отличие от оценки алгоритма AIC, которая обычно оказывается больше реального числа источников помех [11].

С учетом того факта, что сущность обозначенной задачи заключается в разделении множества собственных чисел КМ входного сигнала на два подмножества со значительно отличающимися значениями, представляется перспективным применение при решении данной задачи методов кластерного анализа и обнаружения выбросов в числовых выборках. Один из использовавшихся в рамках данной работы методов этой группы основан на применении в качестве метрики, определяющей включение или невключение элемента выборки в число выбросов, медианного абсолютного отклонения (MAD, *аббр. от англ.* Median Absolute Deviation) числового набора Λ, равного

$$\text{MAD}(\Lambda) = \text{med}(|\Lambda - \text{med}(\Lambda)|),$$

где $\text{med}(\cdot)$ – оператор вычисления медианы выборки.

Известно, что, в отличие от выборочной дисперсии, медианное абсолютное отклонение является робастной оценкой дисперсии конечной выборки [13]. В этой связи можно считать выбросом элемент множества Λ, который отстоит от его медианы более чем на 3MAD(Λ). Положение такого выброса во множестве собственных чисел можно считать местом перехода между двумя подмножествами с сильно отличающимися значениями.

С целью увеличения выраженности перехода между искомыми наборами собственных чисел, уменьшения вариации значений внутри этих наборов и обращения в отрицательные чисел, соответствующих полезным сигналам – обычно близких к нулю – ко множеству собственных значений можно применить логарифмическое преобразование по основанию 10. После этого оценкой количества источников помех можно считать количество положительных значений в получившемся наборе $\{\lg \lambda_j, j = 1, \dots, N\}$.

Наконец, третьим использовавшимся в рамках данного исследования методом разбиения множества собственных чисел на подмножества был известный алгоритм кластеризации по принципу k-средних. Входными данными для данного алгоритма служат числовое множество, подлежащее кластеризации, и требуемое количество кластеров [14]. Это позволяет разделить множество собственных чисел ровно на два подмножества, что приводит к простому решению поставленной задачи.

Алгоритмы определения направлений на источники помех

Все использовавшиеся в рамках данного исследования алгоритмы определения положения источников помех сводятся к задаче поиска экстремумов некоторой функции Q, зависящей от двух переменных – азимута θ и угла места φ. В предположении, что с помощью одного из алгоритмов, описанных в предыдущем подразделе, уже определено число N_I источников помех, среди локальных экстремумов некоторой целевой функции выбираются N_I наиболее значительных и соответствующие им точки экстремума принимаются за направления на источники помех. Приведенные ниже алгоритмы применимы к АР с любым расположением АЭ.

Первый из использовавшихся алгоритмов – метод лучевого сканирования (*от англ.* Beamscan) – в качестве целевой функции имеет следующее выражение:

$$Q_{\text{Beamscan}}(\theta, \varphi) = a^H(\theta, \varphi) R a(\theta, \varphi).$$

За направления на источники помех принимаются N_I локальных максимумов данной функции.

Аналогичным образом устроена целевая функция другого использованного метода – алгоритма Кейпона:

$$Q_{\text{Capon}}(\theta, \varphi) = \frac{1}{a^H(\theta, \varphi) R^{-1} a(\theta, \varphi)}.$$

Так же, как и в предыдущем случае, N_I локальных максимумов данного выражения будут служить оценками направлений на источники помех.

Последним рассмотренным в данной работе методом определения направлений на помехи является известный алгоритм MUSIC, целевую функцию которого можно записать в виде выражения:

$$Q_{\text{MUSIC}}(\theta, \varphi) = a^H(\theta, \varphi) Q_n Q_n^H a(\theta, \varphi),$$

где Q_n – матрица, составленная из собственных векторов КМ входного сигнала, соответствующих

ее $N - N_I$ наименьшим собственным числам. Для отыскания направлений на помехи необходимо найти N_I локальных минимумов указанной функции [11].

Алгоритмы нахождения экстремумов целевых функций для определения направлений на источники помех

Для решения задач оптимизации, возникающих при вычислении направлений на источники помех, необходимо использование алгоритмов поиска экстремумов функции, заданной в прямоугольнике: $\{(\theta, \varphi): 0 \leq \theta < 360, 0 \leq \varphi < 90\}$.

Первый из использовавшихся в данной работе алгоритмов, более простой и один из наиболее распространенных – метод градиентного спуска. Это итерационный алгоритм, на каждом шаге которого вычисляется градиент целевой функции Q в некоторой точке пространства поиска и происходит движение из данной точки либо в направлении полученного градиента (при поиске максимума), либо в противоположном направлении (при поиске минимума) с определенным шагом h. Такое движение должно завершиться в точке локального экстремума при условии подходящего выбора параметров алгоритма – начальной точки движения и величины его шага. Так как ширина пространства поиска в четыре раза превосходит его длину, для успешного нахождения экстремумов целевой функции при количестве помех меньшем или равном трем представляется достаточным разбиение данного прямоугольника на четыре квадрата и выбор каждого из их центров, то есть точек (45,45), (135,45), (225,45) и (315,45) в качестве начальных точек для четырех последовательных запусков алгоритма. Величина шага поиска может быть выбрана равной или кратно большей шага разбиения пространства поиска на равноотстоящие узлы. В качестве критерия достижения локального экстремума и остановки алгоритма можно использовать малость разницы между двумя последовательными значениями целевой функции. К достоинствам данного метода относится вычислительная простота, к недостаткам – неустойчивый характер сходимости в зависимости от величины шага поиска [15].

Другие два алгоритма оптимизации, использованные в данном исследовании, относятся к одним из новейших методов поиска экстремумов функций, основанных на принципах теорий роевого интеллекта и популяционного моделирования, в основе которых лежит интерпретация явлений, имеющих место в живой природе (в частности, среди животных) и их применение для решения различных прикладных задач. Первый из таких методов – алгоритм оптимизации «роем частиц» (*от англ.* Particle Swarm). Перед началом работы данного алгоритма в пространстве поиска случайным образом выбирается некоторое количество точек («частиц»), каждой из которых так же случайно назначается вектор скорости. На каждом шаге алгоритма точки двигаются по пространству поиска в соответствии с уравнениями обновления скорости и положения, а также в зависимости от динамики значений некоторой метрики качества оптимизации, которая заставляет частицы двигаться в направлении лучших позиций относительно их самих и соседних с ними точек. Постепенно большинство частиц достигает оптимального положения, соответствующего некоторому локальному экстремуму оптимизируемой функции.

Похожим образом устроен недавно предложенный алгоритм оптимизации «роем сальп» (*от англ.* Salp Swarm), также использованный в данном исследовании и основанный на моделировании согласованных движений глубоководных морских организмов. Идея данного метода отличается от алгоритма роя частиц движением используемых точек пространства поиска друг за другом, а также выполнением одной итерации в два этапа: исследование пространства поиска и поиск лучших позиций внутри перспективных областей [16].

Преимуществом обоих указанных алгоритмов является отсутствие необходимости знания явного выражения для градиента целевой функции, что значительно упрощает предварительную работу и внутренние вычисления, особенно в тех случаях, когда формулу для градиента оптимизируемой функции слишком сложно или вовсе невозможно выразить относительно одной из независимых переменных. С другой стороны, помимо целевой функции, данные алгоритмы имеют большое число дополнительных входных параметров, которые могут сильно варьироваться в зависимости от особенностей конкретной задачи и нуждаются в надлежащем назначении, зачастую сводящемуся к последовательному ручному подбору [17]. При этом оптимизация «роем сальп», как правило, показывает результаты, сравнимые по качеству с методом «роя частиц», либо лучшие [18, 19].

Для оптимизации «роем частиц» в данной работе использовалась встроенная в среду MATLAB функция «*particleswarm*» (https://www.mathworks.com/help/gads/particleswarm.html). В свою очередь, для оптимизации «роем сальп» была использована ее реализация в MATLAB, выполненная одним из авторов данного метода (https://www.mathworks.com/matlabcentral/fileexchange/63745-ssa-salp-swarm-algorithm).

Компьютерное моделирование пространственной обработки в частотной области

С целью анализа работы всех приведенных выше алгоритмов было проведено компьютерное моделирование ПОС в частотной области в среде MATLAB. Цифровой обработке были подвергнуты записи реальных сигналов от ГНСС с добавлением нескольких пространственно-разнесенных некоррелированных широкополосных энергетических помех. Для приема сигналов использовалась квадратная четырехэлементная АР ($N = 4$). Во время записи сигналов антенная решетка и источники помехи находились в безэховой камере, спутниковый сигнал принимался на крыше здания и подавался в камеру через систему кабелей. Там полезный сигнал и помехи излучались в направлении принимающей АР с помощью передающих антенн, размещенных в разных частях камеры. Антенны, излучавшие помехи, были разнесены друг от друга не менее чем на 90 ° по азимуту с целью обеспечения равномерного заполнения пространства энергетическими помехами. Рассматривались различные сигнально-помеховые ситуации, отличавшиеся количеством источников помех (от одного до трех) и направлением их прихода. В результате проведения данных экспериментов была получена серия многоканальных записей сигналов в виде последовательностей комплексных отсчетов. Среднее значение ОСШ в принятых полезных сигналах при отсутствии помех составило 40 дБ.

При построении ВВК пространственных фильтров по формулам (1–4) используется КМ входных сигналов АР. Однако в реальности ее фактически невозможно вычислить ввиду конечной длительности записей входных сигналов. Поэтому на практике для соответствующих вычислений используются различные выборочные аппроксимации КМ входных сигналов.

При компьютерном моделировании для данной работы в качестве такого приближения была использована матрица, имеющая вид:

$$\hat{R} = \sum_{k=1}^{L_A} x_k x_k^H,$$

где $x_k \in \mathbb{C}^N$ – k-й столбец сигнальной матрицы x; L_A – количество отсчетов сигнала в выборке.

Известно [20], что в случае пространственно-временно́й обработки сигналов оптимальное значение L_A равно $2NN_T$, где N_T – количество временны́х отводов пространственно-временно́го фильтра. При переносе данного результата на случай пространственной обработки в частотной области число N_T будет фактически задавать длину используемого ДПФ. Поэтому для сохранения аналогии с методикой пространственно-временно́й обработки сигналов и обеспечения равенства числа частотных отсчетов T в сегменте обработки степени двойки число L_A было принято равным $16N = 64$.

При моделировании с помощью встроенной в MATLAB функции измерялось время работы каждого из алгоритмов пространственной фильтрации в секундах (ВФ – время фильтрации) с отдельным вычислением времени работы алгоритмов оптимизации (ВО – время оптимизации), а также максимальный поканальный КП помехи в децибелах. Для вычисления прямого и обратного ДПФ использовалась встроенная в MATLAB функция, реализующая быстрое преобразование Фурье. При обработке сигналов алгоритмами, включающими в себя определение числа источников помех и направлений на них, данные показатели также измерялись для последующего сравнения с истинными значениями.

При экспериментах с вычислением количества источников помех были проведены отдельные исследования со всеми возможными случаями расположения источников помех в пространстве. Для определения итоговой оценки числа источников помех вычислялась мода от набора результатов работы всех использовавшихся алгоритмов. При вычислении направлений на помехи шаг сетки поиска был принят равным 1 °. Для определения итоговых направлений на помехи результаты работы всех методов усреднялись. После обработки сигнал подавался на вход программного приемника спутниковых навигационных сигналов SoftGNSS [21], где измерялось среднее ОСШ для спутника, которому соответствует наибольший из максимумов корреляции сигнала с локально генерируемыми опорными С/А-кодами.

Результаты моделирования приведены в таблицах 1–3. Различные алгоритмы оптимизации обозначены в них, соответственно, как ГС (градиентный спуск), РЧ («рой частиц») и РС («рой сальп»).

В таблице 2 в столбце «Положение источников» приняты следующие условные обозначения:
«0» – источник помехи находился на одной высоте с приемником;
«1» – источник помехи находился на 1 м выше приемника;
«–1» – источник помехи находился на 1 м ниже приемника.

Из полученных данных можно сделать следующие выводы. Все рассмотренные алгоритмы фильтрации справляются с поставленной задачей: большие значения КП и ОСШ во всех экспериментах в таблице 1 свидетельствуют о высоком уровне подавления помехи и хорошем качестве полезной составляющей в обработанном сигнале. Наибольшие значения КП и ОСШ во всех экспери-

ментах были получены с использованием алгоритма MPDR. Характеристики подавления алгоритмов LCMP и LCMV примерно равны, однако при малом количестве помех первый метод работает несколько лучше. Значения КП и ОСШ, полученные в результате работы этих алгоритмов, отстоят от соответствующих значений при обработке методом MPDR в среднем на 5–6 дБ. Относительно времени работы лучшие результаты — в среднем на 15 с быстрее остальных — показал алгоритм LCMV. Это обусловлено очень малым временем работы связанных с данным методом алгоритмов оптимизации. Что касается самих алгоритмов поиска экстремума, время их работы оказалось примерно одинаковым. Существенное преимущество алгоритма «роя частиц» заметно лишь в экспериментах с алгоритмом MPDR.

ТАБЛИЦА 1. Результаты экспериментов по моделированию пространственной обработки в частотной области

TABLE 1. Space-Frequency Processing Modeling Experiments Results

N_I	Характеристика работы	Алгоритмы обработки								
		MPDR			LCMP			LCMV		
		ГС	РЧ	РС	ГС	РЧ	РС	ГС	РЧ	РС
1	ВФ, с	18,8	13,2	17,3	21,5	21,1	21,2	9,9	11,9	10,3
	ВО, с	10,7	5,4	9,3	5,5	5,4	5,5	1	1,2	1,1
	КП, дБ	58	58	58	52	52	52	49	49	49
	ОСШ, дБ	45	45	45	41	41	41	37	37	37
2	ВФ, с	18,2	13,9	16,6	21,4	21,3	22	9,8	11	10,4
	ВО, с	9,7	5,7	8	5,2	5,5	5,4	0,8	1,8	1,1
	КП, дБ	63	62	63	56	56	56	50	52	50
	ОСШ, дБ	47	47	47	41	39	41	36	33	34
3	ВФ, с	18,6	10,9	16,7	22,9	21,6	22,1	9,8	9,2	9,7
	ВО, с	10,7	3,1	8,9	7,4	6,4	7	1,2	1	1,2
	КП, дБ	57	57	57	50	50	50	50	50	50
	ОСШ, дБ	41	39	41	35	35	35	35	35	35

Из результатов определения количества помех, приведенных в таблице 2, можно сделать вывод о том, что благодаря использованию моды набора полученных всеми методами результатов, в каждом эксперименте в конечном счете был получен правильный ответ. Однако по отдельности алгоритмы выдавали правильные ответы не в каждом случае. Только методами MDL и MAD во всех случаях были получены правильные ответы. Алгоритм *k*-средних ни разу не выдал неправильный ответ только в случае 2 помех. В остальных случаях наблюдается определенное количество неправильных ответов. В отдельных сериях экспериментов для некоторых алгоритмов количество неправильных ответов превышает половину от числа соответствующих экспериментов – например, в случае с алгоритмом логарифмирования при 3 источниках помех.

ТАБЛИЦА 2. Результаты оценивания количества источников помех в различных экспериментах

TABLE 2. Interference Sources Number Estimation Results in Different Experiments

N_I	Положение источников	AIC	MDL	MAD	lg	*k*-средних	Мода
0	–	0	0	0	4	4	0
1	0	1	1	1	1	4	1
	–1	2	1	1	1	2	1
	1	1	1	1	1	4	1
2	0, 0	3	2	2	2	2	2
	–1, –1	2	2	2	2	2	2
	1, 1	3	2	2	2	2	2
	1, 0	2	2	2	2	2	2
	1, –1	2	2	2	1	2	2
	0, –1	2	2	2	1	2	2
3	0, 0, 0	3	3	3	2	3	3
	–1, –1, –1	3	3	3	3	3	3
	1, 1, 1	4	3	3	3	3	3
	1, 1, 0	4	3	3	3	3	3
	1, 1, –1	3	3	3	2	2	3
	0, 0, –1	3	3	3	2	2	3
	0, 0, 1	4	3	3	2	3	3
	–1, –1, 1	3	3	3	1	3	3
	–1, –1, 0	4	3	3	1	3	3
	1, –1, 0	3	3	3	2	2	3

Из полученных оценок направлений на помехи, приведенных в таблице 3, можно сделать вывод о том, что в целом все алгоритмы определения направлений на источники помех решают поставленную задачу.

ТАБЛИЦА 3. Результаты определения направлений на помехи

TABLE 3. Interference Sources Directions Estimation Results

N_I	Истинные направления	Алгоритмы		
		Beamscan	Capon	MUSIC
1	0 °, 45 °	0 °, 45 °	0 °, 45 °	0 °, 45 °
2	0 °, 45 °	359 °, 41 °	0 °, 45 °	359 °, 42 °
	154 °, 45 °	155 °, 41 °	154 °, 45 °	155 °, 42 °
3	60 °, 45 °	60 °, 32 °	60 °, 44 °	60 °, 29 °
	180 °, 45 °	179 °, 33 °	180 °, 45 °	180 °, 31 °
	300 °, 45 °	301 °, 35 °	300 °, 45 °	301 °, 34 °

Наименее отклоняющиеся от истинных значений результаты показал алгоритм Кейпона. У остальных алгоритмов с увеличением числа помех начинает увеличиваться ошибка определения угла места, хотя азимутальный угол во всех случаях определяется достаточно точно. Самая большая ошибка получилась у алгоритма MUSIC в экспери-

менте с тремя помехами и составила 16 °. Хотя это число и достаточно велико, из успешного подавления помех, которое, как указывалось выше, видно из значений КП и ОСШ в таблице 1, следует, что полученные отклонения от истинных направлений на источники помех недостаточны для существенного ухудшения качества фильтрации.

Заключение

В ходе данного исследования был проведен сравнительный анализ работы нескольких алгоритмов пространственной обработки спутниковых навигационных сигналов в частотной области, повышающей их помехоустойчивость. Компьютерное моделирование сигнальной обработки с использованием записей реальных сигналов от ГНСС показало, что все рассмотренные алгоритмы позволяют успешно подавить энергетические помехи и получить сигнал достаточно хорошего качества. В ходе анализа результатов моделирования были определены алгоритмы с наилучшими показателями работы. Все использованные алгоритмы можно рекомендовать к использованию в реальных устройствах повышения помехоустойчивости спутниковых систем. В зависимости от элементной базы или требований к реализации можно использовать более или менее вычислительно затратные методы.

Список источников

1. Misra P., Enge P. Global Positioning System: Signals, Measurements, and Performance. Ganga-Jamuna Press, 2006. 569 p.

2. Wu R., Wang W., Lu D., Wang L., Jia Q. Adaptive Interference Mitigation in GNSS. Springer, 2018. 274 p. DOI:10.1007/978-981-10-5571-3

3. Gao G.X., Sgammini M., Lu M., Kubo N. Protecting GNSS Receivers From Jamming and Interference // Proceedings of the IEEE. 2016. Vol. 104. Iss. 6. PP. 1327–1338. DOI:10.1109/JPROC.2016.2525938

4. Glushankov E.I., Tsarik V.I. Space-Frequency Beamforming Algorithms Comparison with a Circular Antenna Array // Proceedings of the Conference on Systems of Signals Generating and Processing in the Field of on-Board Communications (Moscow, Russian Federation, 14–16 March 2023). IEEE, 2023. DOI:10.1109/IEEECONF56737.2023.10092000

5. Царик В.И. Сравнение методов определения числа источников помех при адаптации антенных решеток // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2024, Санкт-Петербург, Российская Федерация, 27–28 февраля 2024). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2024. С. 465–469. EDN:OIHERO

6. ГЛОНАСС. Модернизация и перспективы развития. М.: Радиотехника, 2020. 1072 с.

7. ГЛОНАСС. Принципы построения и функционирования. М.: Радиотехника, 2010. 800 с.

8. Малозёмов В.Н., Машарский С.М. Основы дискретного гармонического анализа. СПб.: Лань, 2012. 304 с.

9. Blahut R.E. Fast Algorithms for Signal Processing. New York: Cambridge University Press, 2010. 453 p.

10. Xu H., Cui X., Lu M. An SDR-Based Real-Time Testbed for GNSS Adaptive Array Anti-Jamming Algorithms Accelerated by GPU // Sensors. 2016. Vol. 16. Iss. 356. PP. 1–33. DOI:10.3390/s16030356

11. Van Trees H.L. Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory. New York: John Wiley & Sons, 2002. 1443 p.

12. Пастухов А.В., Оганесян А.А., Головин П.М., Павлов В.С. Мониторинг помеховой обстановки на базе помехоустойчивой адаптивной антенной решётки // Новости навигации. 2015. № 2. С. 8–11. EDN:VLQNYJ

13. Ruppert D., Matteson D.S. Statistics and Data Analysis for Financial Engineering with R examples. New York: Springer, 2015. DOI:10.1007/978-1-4939-2614-5

14. Xu D., Tian Y. A Comprehensive Survey of Clustering Algorithms // Annals of Data Science. 2015. Vol. 2. PP. 165–193. DOI:10.1007/s40745-015-0040-1

15. Аттетков А.В., Галкин С.В., Зарубин В.С. Методы оптимизации. М.: Издательство МГТУ им. Н. Э. Баумана, 2003. 440 с.

16. Mirjalili S., Gandomi A.H., Mirjalili S.Z., Saremi S., Faris H., Mirjalili S.M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems // Advances in Engineering Software. 2017. Vol. 114. PP. 163–191. DOI:10.1016/j.advengsoft.2017.07.002

17. Pedersen M.E.H., Chipperfield A.J. Simplifying Particle Swarm Optimization // Applied Soft Computing. 2010. Vol. 10. Iss. 2. PP. 618–628. DOI:10.1016/j.asoc.2009.08.029

18. Abualigah L., Shehab M., Alshinwan M., Alabool H. Salp swarm algorithm: a comprehensive survey // Neural Computing and Applications. 2020. Vol. 32. PP. 11195–11215. DOI:10.1007/s00521-019-04629-4

19. Houssein E.H., Mohamed I.E., Wazery Y.M. Salp Swarm Algorithm: A Comprehensive Review // Applications of Hybrid Metaheuristic Algorithms for Image Processing. Springer, 2020. PP. 285–308. DOI:10.1007/978-3-030-40977-7_13

20. Reed I.S., Mallett J.D., Brennan L.E. Rapid Convergence Rate in Adaptive Arrays // IEEE Transactions on Aerospace and Electronic Systems. 1974. Vol. AES-10. Iss. 6. PP. 853–863. DOI:10.1109/TAES.1974.307893

21. Borre K., Akos D.M., Bertelsen N., Rinder P., Jensen S.H. A Software-Defined GPS and Galileo Receiver: A Single-Frequency Approach. Boston: Birkhäuser, 2007. 176 p.

References

1. Misra P., Enge P. *Global Positioning System: Signals, Measurements, and Performance*. Ganga-Jamuna Press; 2006. 569 p.
2. Wu R., Wang W., Lu D., Wang L., Jia Q. *Adaptive Interference Mitigation in GNSS*. Springer; 2018. 274 p. DOI:10.1007/978-981-10-5571-3
3. Gao G.X., Sgammini M., Lu M., Kubo N. Protecting GNSS Receivers From Jamming and Interference. *Proceedings of the IEEE*. 2016;104(6):1327–1338. DOI:10.1109/JPROC.2016.2525938
4. Glushankov E.I., Tsarik V.I. Space-Frequency Beamforming Algorithms Comparison with a Circular Antenna Array. *Proceedings of the Conference on Systems of Signals Generating and Processing in the Field of on-Board Communications, 14–16 March 2023, Moscow, Russian Federation*. IEEE; 2023. DOI:10.1109/IEEECONF56737.2023.10092000
5. Tsarik V.I. Comparison of Methods of Interference Sources Number Determination in Antenna Arrays Adaptation. *Proceedings of the VIth International Conference on Infotelecommunications in Science and Education, 27–28 February 2024, St. Petersburg, Russian Federation*. St. Petersburg: The Bonch-Bruevich Saint-Petersburg State University of Telecommunications Publ.; 2024. p.465–469. (in Russ.) EDN:OIHERO
6. *GLONASS. Modernization and development perspectives*. Moscow: Radiotekhnika Publ.; 2020. 1072 p. (in Russ.)
7. *GLONASS. Design and functioning principles*. Moscow: Radiotekhnika Publ.; 2010. 800 p. (in Russ.)
8. Malozyomov V.N., Masharskiy S.M. *Discrete harmonic analysis fundamentals*. St. Petersburg: Lan Publ.; 2010. 304 p. (in Russ.)
9. Blahut R.E. *Fast Algorithms for Signal Processing*. New York: Cambridge University Press; 2010. 453 p.
10. Xu H., Cui X., Lu M. An SDR-Based Real-Time Testbed for GNSS Adaptive Array Anti-Jamming Algorithms Accelerated by GPU. *Sensors*. 2016;16(356):1–33. DOI:10.3390/s16030356
11. Van Trees H.L. *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory*. New York: John Wiley & Sons; 2002. 1443 p.
12. Pastukhov A.V., Oganesyan A.A., Golovin P.M., Pavlov V.S. Monitoring of noise conditions on the basis of an interference-adaptive antenna array. *Novosti navigatsii*. 2015;2:8–11. (in Russ.) EDN:VLQNYJ
13. Ruppert D., Matteson D.S. *Statistics and Data Analysis for Financial Engineering with R examples*. New York: Springer; 2015. DOI:10.1007/978-1-4939-2614-5
14. Xu D., Tian Y. A Comprehensive Survey of Clustering Algorithms. Annals of Data Science. 2015;2:165–193. DOI:10.1007/s40745-015-0040-1
15. Attetkov A.V., Galkin S.V., Zarubin V.S. *Optimization methods: a textbook for universities*. Moscow: Bauman Moscow State Technical University Publ.; 2003. 440 p. (in Russ.)
16. Mirjalili S., Gandomi A.H., Mirjalili S.Z., Saremi S., Faris H., Mirjalili S.M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*. 2017;114:163–191. DOI:10.1016/j.advengsoft.2017.07.002
17. Pedersen M.E.H., Chipperfield A.J. Simplifying Particle Swarm Optimization. *Applied Soft Computing*. 2010;10(2):618–628. DOI:10.1016/j.asoc.2009.08.029
18. Abualigah L., Shehab M., Alshinwan M., Alabool H. Salp swarm algorithm: a comprehensive survey. *Neural Computing and Applications*. 2020;32:11195–11215. DOI:10.1007/s00521-019-04629-4
19. Houssein E.H., Mohamed I.E., Wazery Y.M. Salp Swarm Algorithm: A Comprehensive Review. *Applications of Hybrid Metaheuristic Algorithms for Image Processing*. Springer; 2020. p.285–308. DOI:10.1007/978-3-030-40977-7_13
20. Reed I.S., Mallett J D., Brennan L E. Rapid Convergence Rate in Adaptive Arrays. *IEEE Transactions on Aerospace and Electronic Systems*. 1974;AES-10(6):853–863. DOI:10.1109/TAES.1974.307893
21. Borre K., Akos D.M., Bertelsen N., Rinder P., Jensen S.H. A Software-Defined GPS and Galileo Receiver: A Single-Frequency Approach. Boston: Birkhäuser; 2007. 176 p.

Статья поступила в редакцию 25.09.2024; одобрена после рецензирования 30.10.2024; принята к публикации 18.11.2024.
The article was submitted 25.09.2024; approved after reviewing 30.10.2024; accepted for publication 18.11.2024.

Информация об авторе:

ЦАРИК Владимир Игоревич | ведущий инженер ООО «Эйртэго»
https://orcid.org/0000-0003-3428-9976

Автор сообщает об отсутствии конфликтов интересов.
The author declares no conflicts of interests.

DOI:10.31854/1813-324X-2024-10-6

Научная статья
УДК 004.021
https://doi.org/10.31854/1813-324X-2024-10-6-46-54

Алгоритм обнаружения опорных точек на цифровой электрокардиограмме в режиме реального времени

Белла Кареновна Акопян, akopyan.bella@yandex.ru

Санкт-Петербургский государственный университет аэрокосмического приборостроения,
Санкт-Петербург, 190000, Российская Федерация

Аннотация

Актуальность *темы обусловлена применением цифровых электрокардиографов и кардиомониторов со встроенными алгоритмами автоматической обработки, анализа и интерпретации электрокардиограмм, что позволяет врачу эффективно выполнять диагностику нарушений сердечного ритма. Известно, что для оказания обследуемому экстренной помощи продолжительность диагностики аритмий не должна превышать нескольких десятков секунд, что требует появления новых алгоритмов обнаружения информативных признаков, указывающих на аритмию, работающих в режиме реального времени. Необходимость внедрения новых эффективных технологий диагностики сердечно-сосудистых заболеваний также отражена в государственных программах развития здравоохранения.*

Целью *исследования является разработка и анализ показателей качества алгоритма обнаружения опорных точек на цифровой электрокардиограмме, несущих информативные признаки для процедуры диагностики аритмий.*

Используемые методы. *Исследование основано на анализе существующих подходов к решению задачи обнаружения опорных точек на цифровой электрокардиограмме, а также проведении экспериментальной проверки предлагаемого алгоритма методами математического моделирования. Предложены показатели качества рассматриваемых алгоритмов, определенные в соответствии с принципами теории обнаружения сигналов и диагностического тестирования, на стыке которых расположена задача обнаружения опорной точки кардиокомплекса. Экспериментальная проверка предлагаемого алгоритма осуществлена на материалах открытой верифицированной базы данных MIT-BIH Arrhythmia Database, которая широко применяется для верификации и валидации алгоритмов обработки сигнала цифровой электрокардиограммы, работающих в режиме реального времени.*

Решение. *В работе предложен алгоритм обнаружения опорных точек на цифровой электрокардиограмме, который основан на цифровой фильтрации сигнала с применением решающего правила на базе трехэтапной двухпороговой схемы сравнения величин сигнала предобработанной электрокардиограммы на скользящем окне, обладающий элементами* ***научной новизны****. Эксперимент на материалах открытой верифицированной базы данных MIT-BIH Arrhythmia Database показал, что качество предложенного алгоритма обнаружения опорных точек выше, чем у алгоритмов, применяемых в современных цифровых электрокардиографах и кардиомониторах.*

Значимость. *Полученные в работе результаты могут быть использованы при разработке устройств цифрового мониторинга сердечно-сосудистой системы, а также для автоматической обработки, анализа и интерпретации сигнала цифровой электрокардиограммы в режиме реального времени с применением ЭВМ.*

Ключевые слова: *цифровая электрокардиография, электрокардиограмма, аритмия, информативный признак, опорные точки, алгоритм, диагностика*

Ссылка для цитирования: Акопян Б.К. Алгоритм обнаружения опорных точек на цифровой электрокардиограмме в режиме реального времени // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 46–54. DOI:10.31854/1813-324X-2024-10-6-46-54. EDN:ADHKYB

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-46-54

Algorithm for Detecting Reference Points on a Digital Electrocardiogram in Real Time

Bella K. Akopyan, akopyan.bella@yandex.ru

Saint Petersburg State University of Aerospace Instrumentation,
St. Petersburg, 190000, Russian Federation

Annotation

Relevance*. The use of digital electrocardiographs and cardiac monitors with built-in algorithms for automatic processing, analysis and interpretation of electrocardiograms allows the doctor to effectively diagnose cardiac arrhythmias. It is known that in order to provide emergency care to a patient, the duration of arrhythmia diagnostics should not exceed several tens of seconds, which requires the emergence of new algorithms for detecting informative features indicating arrhythmia, operating in real time. The need to introduce new and effective technologies for diagnosing cardiovascular diseases is also reflected in public health development programmes.*

Research goal. *Development and quality indicators analysis of the algorithm for reference points detection on a digital electrocardiogram, bearing informative signs for the procedure of arteries diagnosis.*

The methods used*. The study is based on an analysis of existing approaches to the problem of reference points detection on digital electrocardiogram, as well as conducting a test of the proposed algorithms by mathematical modelling methods. The quality indicators of the algorithms defined in accordance with the principles of signal detection theory and diagnostic testing, at the junction of which the task of electrocardiogram reference point detection is located. The proposed algorithm was tested on materials of MIT-BIH Arrhythmia Database, which is widely used for verification and validation of real-time digital electrocardiogram signal processing algorithms.*

The results. *The study proposes an algorithm for detecting reference points on a digital electrocardiogram that carry informative features for the arrhythmia diagnostic procedure. The proposed algorithm is based on digital signal filtering using a decision rule based on a three-step two-threshold principle of pre-processed electrocardiogram signal values comparison on a sliding window. An experiment on the materials of the open verified MIT-BIH Arrhythmia DB showed that the quality of the proposed algorithm for detecting reference points is higher than that of the algorithms used in modern digital electrocardiographs and cardiac monitors. The proposed algorithm based on digital signal filtering and the three-step two-threshold decision rule have elements of* ***scientific novelty****.*

The significance*. The results of this work can be used in the development of digital heart rate monitors, cardiac devices and for automatic processing, analysis and real-time computer-assisted digital electrocardiogram signal interpretation.*

Keywords: *digital electrocardiography, electrocardiogram, arrhythmia, informative sign, reference points, algorithm, diagnostics*

For citation: Akopyan B.K. Algorithm for Detecting Reference Points on a Digital Electrocardiogram in Real Time. *Proceedings of Telecommunication Universities.* 2024;10(6):46–54. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-46-54. EDN:ADHKYB

Введение

Электрокардиография относится к самым распространенным методам инструментального обследования. Она широко применяется при кардиологических исследованиях за счет своей информативности, простоты, доступности и безопасности, а также невысокой себестоимости [1].

Современная электрокардиография характеризуется широким применением цифровых электрокардиографов и кардиомониторов со встроенными алгоритмами автоматической обработки, анализа и интерпретации электрокардиограмм (ЭКГ). В частности, такие устройства широко применяются для диагностики нарушений сердечного ритма. В целях получения достоверной диагностики в

настоящее время приходится прибегать к длительной регистрации ЭКГ и отсроченному анализу ранее записанных фрагментов ЭКГ. Но известно, что для оказания пациенту экстренной помощи продолжительность диагностики аритмий не должна превышать нескольких десятков секунд [1], в связи с чем целесообразно разработать алгоритмы обнаружения аритмий, работающие в режиме реального времени.

Электрокардиосигнал (ЭКС) – сигнал, который несет информацию об изменениях во времени суммарного электрического потенциала, возникающего в сердечной мышце за счет движения ионов через мышечную мембрану. Один период сокращений сердечной мышцы принято называть сердечным циклом или кардиоциклом (КЦ).

ЭКГ представляет собой запись отсчетов ЭКС. Схематическое изображение кривой ЭКГ нормальной формы представлено на рисунке 1. За нулевой уровень ЭКС принимается изоэлектрическая линия (изолиния) – горизонтальная прямая, указывающая на отсутствие электрической активности [2]. Отклонение от нее указывает на электрическую активность сердечных мышц.

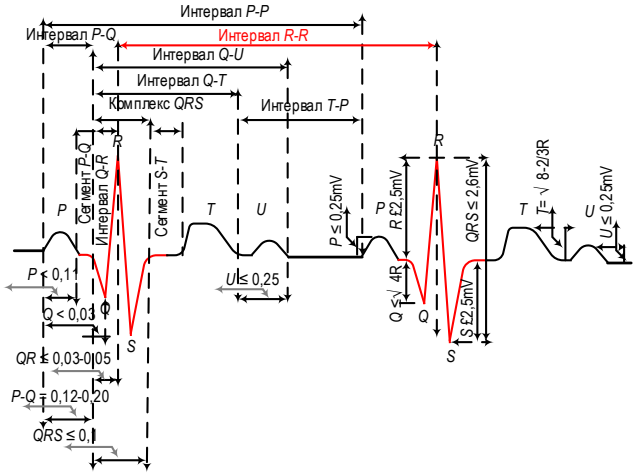


Рис. 1. Схематическое изображение электрокардиограммы нормальной формы с отображением информативных фрагментов [2]

Fig. 1. Schematic Image of Normal-Form Electrocardiogram (ECG) with Informative Fragments [2]

Каждый отдельный КЦ представлен на ЭКГ функцией сложной формы. Информативные фрагменты кривой КЦ отражают стадии прохождения волны возбуждения по отдельным участкам сердца. К информативным сегментам ЭКГ относятся зубцы, сегменты и интервалы [2].

Зубцы отражают работу определенных участков сердца и внешне представляют собой набор пиков и впадин. Их разделяют на отрицательные (расположенные ниже изолинии) и положительные (выше изолинии). Зубцы ЭКГ стандартно обозначаются латинскими буквами слева направо в порядке своего появления. В ЭКГ всегда выделяются зубцы *P, Q, R, S, T,* в некоторых случаях выделяется низкоамплитудный зубец *U,* следующий за *T*. Зубец *P* отображает процесс деполяризации миокарда предсердий, зубцы *Q, R* и *S* – деполяризации желудочков (объединяются в желудочковый *QRS*-комплекс); зубец *T* связан с реполяризацией миокарда желудочков [2].

Сегмент – отрезок изоэлектрической линии, заключенный между двумя соседними зубцами; в случае ЭКГ нормальной формы он не искажен и не смещен относительно изолинии. Среди сегментов ЭКГ большое внимание при диагностике уделяется сегменту *ST* – отрезку кривой ЭКГ между концом *QRS*-комплекса и началом зубца *T*, который соответствует периоду сердечного цикла, когда оба желудочка полностью охвачены возбуждением. Искажения этого информативного сегмента характерны для многих сердечно-сосудистых заболеваний (инфаркт миокарда, ишемическая болезнь сердца и др.).

Интервал, в свою очередь, состоит из зубца (или комплекса зубцов) и сегмента (https://www.biors.ru/tech/practicing-biors/konturniy-analiz-ekg.htm). Для оценки сердечного ритма традиционно применяется измерение *RR*-интервалов, поскольку данный параметр достаточно точно характеризует состояние ритма и отличаются высокой надежностью при измерениях в условиях различных помех. В таблице 1 приведены значения амплитудно-временны́х параметров нормальной ЭКГ.

ТАБЛИЦА 1. Параметры информативных фрагментов нормальной ЭКГ [3]

TABLE 1. Parameters of Normal ECG Informative Fragments [3]

Параметр	Информативный фрагмент			
	Зубец *P*	*QRS*-комплекс	Сегмент *ST*	Зубец *T*
Амплитуда, мВ	0–0,25	0,3–5	–0,1–0,1	0,4–1
Длительность, с	0,07–0,11	0,06–0.1	0,06–0,15	0,1–0,2

Анализ сердечного ритма и его возможные нарушения

Анализ сердечного ритма включает в себя оценку регулярности сердечных сокращений и подсчет их числа за единицу времени. Принято считать, что ритм правильный, если продолжительность *RR*-интервалов при постоянном физиологическом состоянии обследуемого отличается не более, чем на ±10 % [3].

Показателем числа сердечных сокращений ЭКС является их частота (ЧСС), измеряемая числом ударов в минуту:

$$F_{\text{ЧСС}} = \frac{60}{t_{RR}}, \qquad (1)$$

где t_{RR} – длительность *RR*-интервала в секундах.

Принято считать, что нормальная ЧСС составляет 60–90 ударов в минуту. Состояние при ЧСС выше 90 ударов в минуту – тахикардия, понижение ниже, чем 60 – брадикардия (критическая ЧСС составляет 30–35 ударов в минуту) [3]. В случае, если ритм сердечных сокращений не соответствует нормальному вследствие нарушений в работе сердца, принято говорить о сердечной аритмии. К ней относятся экстрасистолия, мерцательная аритмия, синусовая аритмия и т. д.

Экстрасистолия – наиболее распространенный вид аритмий, представляющий собой явление преждевременного внеочередного возбуждения сердца; она может встречаться даже у здоровых людей вследствие чрезмерных физических и эмоциональных нагрузок. Экстрасистолы являются симптомом таких болезней, как ишемическая болезнь сердца, воспаление сердечной мышцы, сердечная недостаточность, дефект сердечного клапана и т. д. (https://empendium.com/ru/chapter/B33.II.2.6.1).

Вид экстрасистол непосредственно связан с областью локализации. В зависимости от этого фактора выделяют предсердные, желудочковые и атриовентрикулярные (предсердно-желудочковые) экстрасистолы. При анализе экстрасистол оперируют понятиями «интервал сцепления» и «компенсаторная пауза». Интервал сцепления – расстояние от КЦ основного ритма, предшествующего экстрасистоле, до самой экстрасистолы. Компенсаторная пауза – расстояние от экстрасистолы до следующего за ней кардиоцикла основного ритма. Если в сумме интервал сцепления и компенсаторная пауза дают два *RR*-интервала, то говорят о полной компенсаторной паузе; если сумма меньше, то – о неполной. Требуется отметить, что для предсердных экстрасистол при измерениях интервала сцепления и компенсаторной паузы вместо *RR*-интервалов может применяться *PP*-интервал, но, поскольку при таких экстрасистолах существует вероятность наслоения зубца *P* на *QRS*-комплекс и *T*-зубец, это не всегда целесообразно (https://empendium.com/ru/chapter/B33.II.2.6.1).

Характерными признаками – ЭКГ-критериями – предсердных экстрасистол являются:

– преждевременное внеочередное появление зубца *P,* который также может сменить полярность или оказаться деформированным;

– следующий за преждевременным зубцом *P* желудочковый *QRS*-комплекс не претерпел изменений;

– компенсаторная пауза неполная.

Наиболее существенными ЭКГ-критериями желудочковой экстрасистолии является преждевременное внеочередное появление расширенного и деформированного *QRS*-комплекса, возможно отсутствие зубца *P*, в большинстве случаев компенсаторная пауза полная (https://empendium.com/ru/chapter/B33.II.2.6.1).

Отметим, что появление на ЭКГ экстрасистол может свидетельствовать о вероятности возникновения других, более опасных для жизни аритмий сердца (https://compendium.com.ua/clinical-guidelines/cardiology/section-13/glava-1-diagnostika-i-lechenie-ekstrasistolii-i-parasistolii). В частности, распространенным примером является мерцательная аритмия (фибрилляция предсердий) – состояние, при котором наблюдается частое беспорядочное возбуждение и сокращение отдельных групп мышечных волокон предсердий. Предвестниками мерцательной аритмии часто являются множественные экстрасистолы, в связи с чем при классификации аритмии допускается относить множественную экстрасистолию к фибрилляции предсердий [4].

На рисунке 2 представлена ЭКГ с мерцательной аритмией в сравнении с ЭКГ нормальной формы.

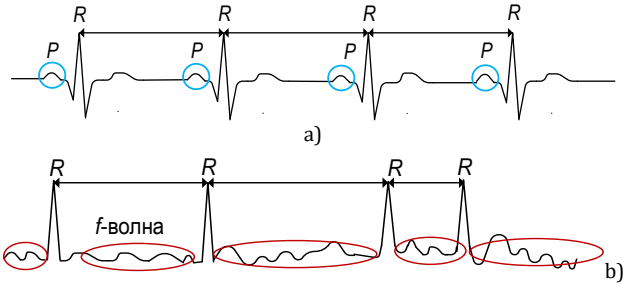


Рис. 2. ЭКГ: a) нормальной формы; b) с мерцательной аритмией [5]

Fig. 2. EKG: a) Normal; b) with Atrial Fibrillation [5]

ЭКГ-критериями мерцательной аритмии являются отсутствие зубца *P*, наличие *QRS*-комплексов без деформаций и уширения, нерегулярный ритм, неодинаковые *RR*-интервалы, наличие на протяжении всего сердечного цикла мелких беспорядочных *f*-волн [6].

Общие сведения об алгоритмах обнаружения аритмий

Обобщенный принцип работы алгоритмов классификации аритмических эпизодов по сигналу ЭКГ состоит из следующих этапов.

1. Предобработка сигнала ЭКГ.
2. Обнаружение опорных точек ЭКГ.
3. Оценивание ЭКГ по информативным признакам.
4. Формирование заключения.

После приема сигнала необходима его первичная обработка (предобработка) – установка факта наличия на анализируемом участке КЦ с последующим определением, является ли ритм нормальным или имеет место патология. Стоит отметить, что поскольку термин «аритмия» объединяет различные по механизму, клиническим проявлениям и прогностическому значению нарушения [5], то алгоритмы обнаружения аритмий могут не только регистрировать возможные нарушения сердечного

ритма, но и осуществлять простейшую классификацию аритмий по ЭКГ-признакам патологий.

Обнаружение опорных точек ЭКГ является очень важным этапом работы алгоритмов обнаружения аритмий, поскольку от качества его функционирования напрямую зависит точность и достоверность дальнейшей диагностики состояния ритма и возможной классификации [7, 8]. Опорные точки определяются на этапе предобработки сигнала ЭКГ. Так, для измерения длительности *RR*-интервалов в качестве опорной точки целесообразно принять точку максимума *R*-зубца.

Этап оценивания ЭКГ по информативным признакам включает как обнаружение самих информативных признаков, так и длительностей интервалов – и те, и другие показаны на рисунке 1. Выявление информативных признаков, как правило, реализуется методами обнаружения *QRS*-комплексов. Общей чертой указанных методов является вид решающего правила: отсчет выборки ЭКГ после специальной обработки сравнивается с неким пороговым значением, и если порог превышен, то возможно обнаружен *QRS*-комплекс – так называемый *QRS*-кандидат.

Чаще всего в качестве преобразованной выборки применяются [9–11]:

– массив производных записи ЭКГ: фрагмент выборки ЭКГ принимается в качестве *QRS*-кандидата, когда некоторое количество последовательных отсчетов массива производных превышают пороговое значение (обычно применяются производные не выше второго порядка);

– выборка ЭКГ, прошедшая цифровую фильтрацию;

– выборка ЭКГ, над которой осуществлено преобразование формы / масштаба (преобразование Гильберта, вейвлет-преобразование и др.);

– массив коэффициентов корреляции ЭКГ: фрагмент выборки принимается в качестве *QRS*-кандидата, когда коэффициент корреляции между значениями эталонного *QRS*-комплекса и выборки ЭКГ превышает пороговое значение (http://www.vestar.ru/atts/10480/HRV%20standards.pdf); в качестве эталона *QRS*-комплекса можно использовать как предварительно заданные функции, так и один из *QRS*-комплексов обследуемого, записанный в начале работы регистрирующего устройства.

Анализ сердечного ритма, как правило, выполняется методами *RR*-интервалов, анализ которых позволяет определить текущую ЧСС и сформировать заключение об основных характеристиках ритма. Очевидно, что результаты принятия решения о нарушении ритма напрямую зависят от результатов обнаружения опорных точек, поскольку по обнаруженным *R*-зубцам рассчитываются значения *RR*-интервала и на их основании формируются предположения о характере сердечного ритма.

Методы анализа нарушений сердечного ритма во многом близки с методами оценки вариабельности. Но, в отличие от показателей вариабельности, которые могут вычисляться по истечении определенного периода, нарушения ритма должны фиксироваться в режиме реального времени. Методы можно разделить на две группы (http://www.vestar.ru/atts/10480/HRV%20standards.pdf):

1) полученные при обработке прямых измерений длительности *RR*-интервалов;

2) вычисленные на основе разницы между *RR*-интервалами.

При выборе алгоритмов обнаружения и классификации аритмий принято опираться на вероятности ошибок первого и второго рода [12, 13], причем гипотезы на каждом из промежуточных этапов обработки ЭКГ различаются. На этапе обнаружения опорных точек ЭКГ ошибкой первого рода является ложное обнаружение опорной точки на интервале, где ее объективно нет, ошибкой второго рода – пропуск истинной опорной точки. На этапе анализа ритма ЭКГ ошибкой первого рода является принятие решения о нарушении ритма, когда он нормальный, ошибкой второго рода – принятие решения о нормальном состоянии ритма в случае, когда объективно имеет место его нарушение.

Поскольку и ложнообнаруженные *R*-зубцы, и пропуски истинных *QRS*-комплексов в равной степени могут повлиять на результаты оценки длительности *RR*-интервала и последующее принятие решения о нарушении ритма, то в качестве основного критерия отбора подходящего алгоритма обнаружения *QRS*-комплексов целесообразно использовать вероятности ошибки и первого, и второго рода.

Аналогично, на этапе анализа ритма и ложное решение о его нарушении, и ложное решение о его нормальном состоянии являются в равной степени неприемлемыми (первое – из-за дискредитации самой идеи автоматической диагностики, второе – из-за опасности оставления больного без адекватного реагирования).

Таким образом, исходя из практических соображений для реализации в носимом устройстве диагностики, к разрабатываемым и исследуемым алгоритмам предъявлены следующие требования:

– реализуемость в режиме реального времени;

– нечувствительность к низким уровням сигнала аддитивной помехи ЭКГ во избежание ложных срабатываний в реальных условиях;

– короткий период предварительной настройки алгоритма для его адаптации к конкретным условиям.

Описание предложенного алгоритма обнаружения опорных точек ЭКГ

Блок-схема нового алгоритма обнаружения опорных точек ЭКГ приведена на рисунке 3.

Прошедший предобработку сигнал ЭКГ пропускается через специальное дифференцирующее устройство, математическая модель которого описывается выражением:

$$Y0_n = s_n - s_{n-4}, n = 4{,}5 \dots L - 1, \quad (1)$$

где $Y0$ – сигнал на выходе дифференциатора; s – исходный ЭКС; L – объем выборки скользящего окна.

Данная процедура нейтрализует остаточное изменение уровня сигнала после предобработки. Затем полученные данные пропускаются через цифровой нерекурсивный фильтр нижних частот (ФНЧ) с частотой среза 62,5 Гц, обусловленной спектральными характеристиками кардиокомплекса.

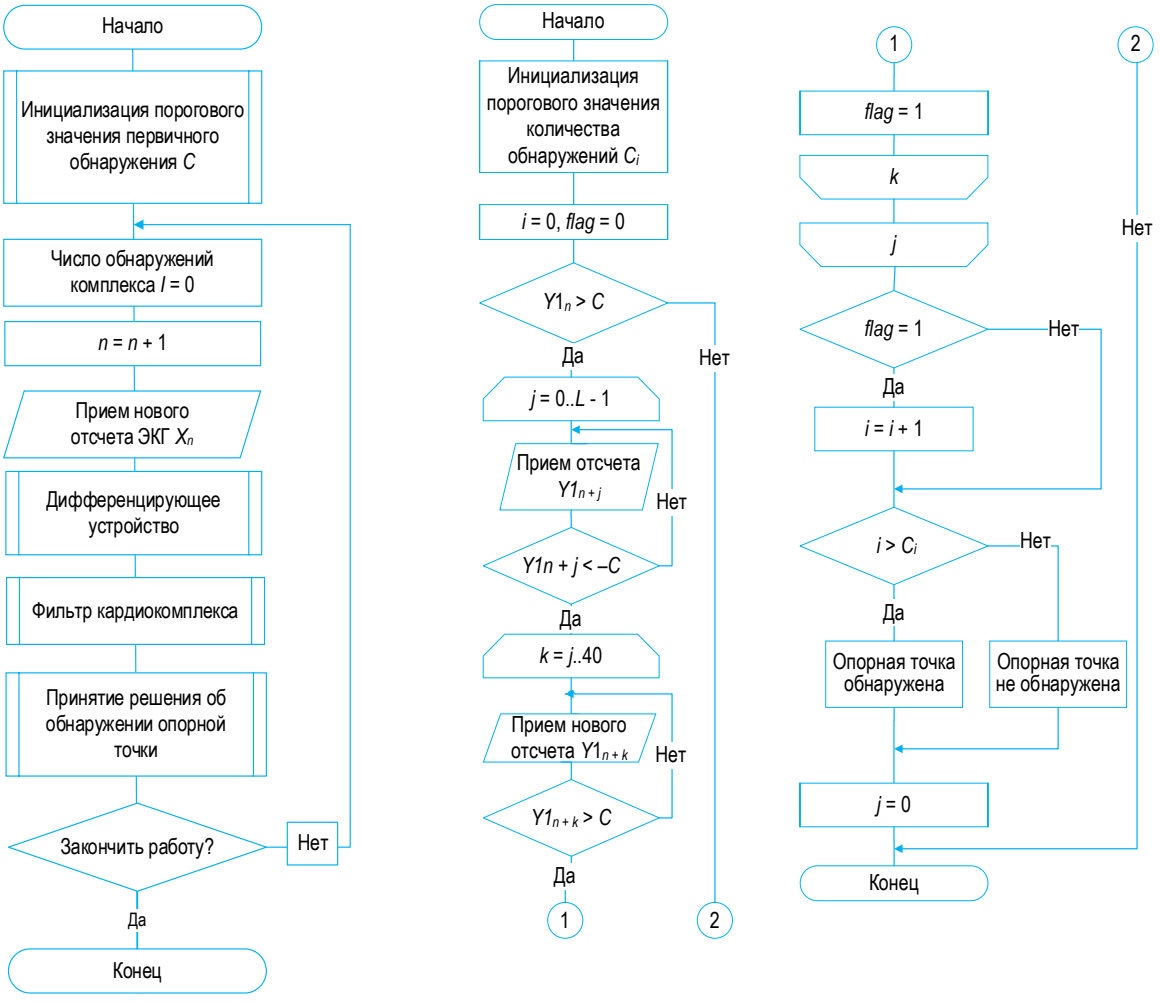
Математическая модель ФНЧ описывается следующим выражением:

$$Y1_n = \sum_{i=0}^{4} p_i y0_{n-i}, p = \{1,4,6,4,1\}, \quad (2)$$

где $Y1$ – сигнал на выходе ФНЧ; p – массив коэффициентов фильтра.

Полученный сигнал проходит через трехэтапную двухпороговую схему сравнения: пороговые значения C равны по величине, но противоположны по знаку. Пороговое значение C определяется адаптивно в течение первых 5 секунд после начала измерения и рассчитывается в соответствии с обеспечивающим лучшие характеристики в ходе вычислительных экспериментов выражением:

$$C = \frac{5}{3} \cdot \left(\frac{1}{L} \sum_{i=0}^{L-1} y1_i \right)^2. \quad (3)$$



a) b)

Рис. 3. Блок-схема предлагаемого алгоритма обнаружения опорной точки ЭКГ: a) алгоритм обнаружения опорной точки; b) решающее правило

Fig. 3. Flowchart of the Proposed ECG Reference Point Detection Algorithm: a) Reference Point Detection Algorithm; b) Decision Algorithm

Сигнал на выходе ФНЧ сканируется до тех пор, пока не будет обнаружен отсчет, величина которого больше положительного порогового значения, что является началом области поиска длительностью в L отсчетов. В этом случае, опорную точку можно считать правильно обнаруженной, если определенный алгоритмом отсчет будет отклоняться от истинного положения не более, чем на $\pm L/2$ отсчетов.

Наличие дополнительных пересечений порога используется для классификации результата обнаружения как потенциальной опорной точки (R-кандидата) или вызванное остаточными артефактами сигнала помехи. Если в течение L последующих отсчетов не происходит ни одного пересечения порогового значения, то превышение порога было спровоцировано дрейфом изолинии и точка исключается из рассмотрения. В противном случае по очереди проверяются условия:

$$Y1_{n+j} < -\mathrm{C},\ Y1_{n+k} > \mathrm{C},\ 0 < j < L-1,\quad j < k < L-1. \tag{4}$$

Если все условия выполняются, то определена потенциальная опорная точка и регистрируется количество повторных обнаружений данной опорной точки. Для них введено дополнительное пороговое значение: если количество обнаружений одного R-кандидата i превышает пороговое значение C_i, распознанная опорная точка фиксируется, в противном случае исключается из рассмотрения.

Поскольку и ложное обнаружение R-зубца, и пропуск истинного кардиокомплекса в равной степени приводят к одинаково опасным последствиям принятия решения о нарушении ритма, то в качестве критерия отбора подходящего алгоритма обнаружения опорных точек используются вероятности ошибки первого и второго рода (p_1 и p_2):

$$p_1 = \frac{FP}{TP+FN},\ p_2 = \frac{FN}{TP+FN}, \tag{5}$$

где TP – число истинно-положительных решений; FP – число ложноположительных решений; FN – число ложноотрицательных решений.

Определено, что, поскольку во множестве результатов обнаружения опорных точек отсутствует истинно-отрицательный результат, то из показателей качества диагностических тестов целесообразно унаследовать показатели чувствительности Sn и предиктивности PV, которые отражают предсказательную ценность положительного результата, а именно долю истинно-положительных результатов относительно всех кардиокомплексов и всех положительных результатов, полученных алгоритмом:

$$Sn = \frac{TP}{TP+FN},\ PV = \frac{TP}{TP+FP}. \tag{6}$$

Вычисленные параметры соответствуют показателям полноты и точности алгоритма, поэтому их можно объединить в единый показатель, аналогичный сбалансированной F-мере:

$$F_{QRS} = 2\frac{Sn \cdot PV}{Sn+PV}. \tag{7}$$

В работе проведена оптимизация значений параметров порогового значения числа обнаружений C_i и размера скользящего окна L, которые обеспечили наилучшее выделение опорной точки кардиокомплекса по заданным показателям качества. На рисунке 4 приведены результаты математического моделирования процедуры обнаружения опорных точек разработанным алгоритмом при оптимальных значениях параметров C_i = 10, L = 40, где 1 – запись ЭКГ, 2 – истинные R-зубцы, 3 – обнаруженные R- зубцы, 4 – экстрасистолы.

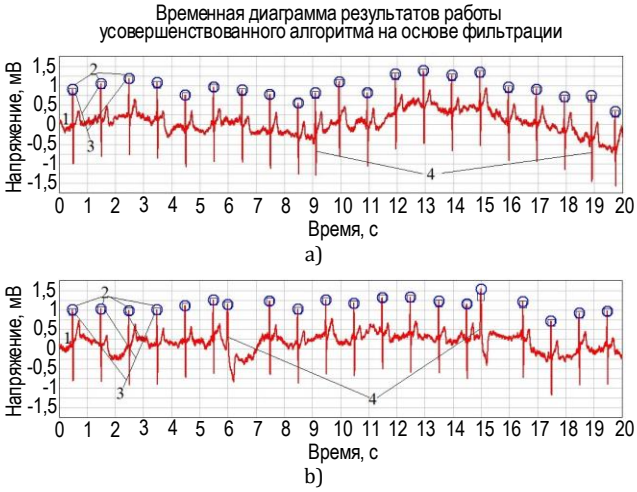


Рис. 4. График результатов обнаружения опорных точек кардиокомплексов на ЭКГ алгоритмом на основе цифровой фильтрации на промежутке [0;20] с: a) запись с предсердной экстрасистолией; b) запись с желудочковой экстрасистолией

Fig. 4. Timing Diagram of the Cardiac Complexes Detection Results by ECG Algorithm Based on Digital Filtration at Interval [0;20] s: a) Atrial Extrasystole ECG Record; b) Ventricular Extrasystole ECG Record

Для оценки эффективности разработанного алгоритма были выбраны процедуры, применяемые для обнаружения опорных точек в режиме реального времени и соответствующие приведенным ранее практическим соображениям: алгоритм Пана – Томпкинса и корреляционно-экстремальный алгоритм. Исследование проводилось на материалах открытой верифицированной базы данных MIT-BIH Arrhythmia Database. Результаты сравнительного анализа показателей качества приведены в таблице 2. Очевидно, что предложенный алгоритм демонстрирует высокие показатели качества по сравнению с иными существующими алгоритмами.

ТАБЛИЦА 2. Показатели качества алгоритмов обнаружения опорных точек, рассматриваемых в исследовании

TABLE 2. Quality Indicators of the Algorithms for Reference Points Detection Considered in the Study

Алгоритм	P	σ_P	F_{QRS}, %
Корреляционно-экстремальный	0,0406	0,0106	97,48
Пан-Томпкинс	0,0041	0,0012	99,73
Алгоритм фильтрации, $C_i = 10$, $L = 40$	0,0012	0,0004	99,92

Таким образом, предложенный алгоритм обнаружения опорных точек ЭКГ демонстрирует эффективность для кардиокомплексов, результат работы которого является основой для разработки методов выявления нарушений сердечного ритма.

Заключение

В работе обсуждается задача разработки алгоритмического обеспечения цифровых электрокардиографов и кардиомониторов, позволяющих в реальном времени диагностировать нарушения сердечного ритма. Представлены общие принципы работы алгоритмов обнаружения аритмий по ЭКГ. Рассмотрены информативные признаки, регистрирование которых позволяет обнаружить возможные нарушения сердечного ритма и осуществлять простейшую классификацию аритмий по ЭКГ-признакам патологий. Показано, что обнаружение опорных точек ЭКГ является очень важным этапом работы алгоритмов обнаружения аритмий, поскольку от качества его работы напрямую зависит точность и достоверность дальнейшей диагностики состояния ритма и возможной классификации. Предложен новый алгоритм обнаружения опорных точек, основанный на цифровой фильтрации *QRS*-комплексов с применением решающего правила на базе трехэтапной двухпороговой схемы сравнения величин сигнала предобработанной электрокардиограммы на скользящем окне. Показатели качества предложенного алгоритма обнаружения опорных точек выше, чем у корреляционно-экстремального алгоритма и алгоритма Пана – Томпкинса, применяемых в современных цифровых электрокардиографах и кардиомониторах.

Список источников

1. Юлдашев З.М. Продолжительность диагностики аритмий для оказания экстренной помощи не должна превышать нескольких десятков секунд // Медвестник. 2018. URL: https://medvestnik.ru/content/interviews/Prodoljitelnost-diagnostiki-aritmii-dlya-okazaniya-ekstrennoi-pomoshi-ne-doljna-prevyshat-neskolkih-desyatkov-sekund.html (дата обращения 30.09.2024)

2. Нестерова Е.А. Основы электрокардиографии. Нормальная ЭКГ // Кардиология: Новости. Мнения. Обучение. 2016. № 2(9). С. 77–85. EDN:WFLXIP

3. Рудницкий Л.В. Карманный справочник медицинских анализов. СПб.: Питер, 2014. 320 с.

4. Акопян Б.К. Классификация эпизодов нарушений сердечного ритма по информативным признакам во временной области электрокардиограммы // Известия высших учебных заведений. Приборостроение. 2024. Т. 67. № 4. С. 305–314. DOI:10.17586/0021-3454-2024-67-4-305-314. EDN:DSWAXC

5. Иванов Г.Г., Дворников В.Е., Сбеитан С., Булгакова Е.Ю., Александрова М.Р., Грибанов А.Н. Анализ показателей структуры вариабельности ритма сердца у здоровых лиц по данным PP- и RR-интервалов // Вестник Российского университета дружбы народов. Серия: Медицина. 2007. № 4. С. 26–34. EDN:JVDXGH

6. Анциперов В.Е., Забросаев И.В., Растягаев Д.В. Детектирование нарушений сердечного ритма с использованием техники аналитических спектров // Журнал радиоэлектроники. 2015. № 12. С. 16. EDN:VHTMVJ

7. Friesen G.M., Jannett T.C., Jadallah M.A., Yates S.L., Quint S.R., Nagle H.T. A comparison of the noise sensitivity of nine QRS-detection algorithms // IEEE Transactions on Biomedical Engineering. 1990. Vol. 37. Iss. 1. PP. 85–98. DOI:10.1109/10.43620

8. Akopyan B. Development of the automated cardiac rhythm disorders detection and classification algorithm // Bulletin of the UNESCO Department “Distance education in engineering” of the SUAI. 2022. Iss. 7. PP. 28–31. EDN:QSKIML

9. Kohler B.-U., Hennig C., Orglmeister R. The principles of software QRS detection // IEEE Engineering in Medicine and Biology Magazine. 2002. Vol. 21. Iss. 1. PP. 42–57. DOI:10.1109/51.993193

10. Zong W., Moody G.B., Jiang D. A robust open-source algorithm to detect onset and duration of QRS-complexes // Proceedings of the Conference on Computers in Cardiology (Thessaloniki, Greece, 21–24 September 2003). IEEE, 2003. PP. 737–740. DOI:10.1109/CIC.2003.1291261

11. Жаринов О.О., Жаринов И.О. Применение корреляционно-экстремального метода для решения задач обнаружения и оценивания положений опорных точек QRS-комплексов в электрокардиограмме // Научно-технический вестник информационных технологий, механики и оптики. 2011. № 5(75). С. 85–90. EDN:OCBFFH

12. Боженко В.В., Черныш Н.Ю., Татарникова Т.М. Интеллектуальный анализ данных в диагностике анемии по клиническим показателям // Известия высших учебных заведений. Приборостроение. 2024. Т. 67. № 4. С. 321–329. DOI:10.17586/0021-3454-2024-67-4-321-329. EDN:AUAHNY

13. Раскопина А.С., Боженко В.В., Татарникова Т.М. Использование глубокого обучения при диагностировании пневмонии по рентгеновским снимкам // Известия высших учебных заведений. Приборостроение. 2024. Т. 67. № 4. С. 315–320. DOI:10.17586/0021-3454-2024-67-4-321-329. EDN:UPSNQQ

References

1. Yuldashev Z.M. The duration of arrhythmia emergency diagnosis should not exceed several tens of seconds. *Medvestnik*. 2018. (in Russ.) URL: https://medvestnik.ru/content/interviews/Prodoljitelnost-diagnostiki-aritmii-dlya-okazaniya-ekstrennoi-pomoshi-ne-doljna-prevyshat-neskolkih-desyatkov-sekund.html [Accessed 30.09.2024]
2. Nesterova E.A. Electrocardiography Bases. Normal Electrocardiography (The Module for Continuous Medical Education). *Cardiology: News. Opinions. Training*. 2016;2(9):77–85. (in Russ.) EDN:WFLXIP
3. Rudnitskiy L.V. *Pocket Guide to Medical Analysis*. St. Petersburg: Piter Publ.; 2014. 320 p. (in Russ.)
4. Akopyan B.K. Classification of heart rhythm disorder episodes by informative features in the electrocardiogram time domain. *Journal of Instrument Engineering*. 2024;67(4):305–314. (in Russ.) DOI:10.17586/0021-3454-2024-67-4-305-314. EDN:DSWAXC
5. Ivanov G.G., Dvornikov V.E., Sbeytan S., Bulgakova E.Ju., Alexandrova M.R., Gribanov A.N. The Analysis of Parameters Structure Variability of Heart Rhythm at Healthy Persons on Data PP-and RR-Intervals Determinants of Evolution of ECG Signs of Left Ventricular. *RUDN Journal of Medicine.* 2007;4:26–34. (in Russ.) EDN:JVDXGH
6. Antsiperov V.E., Zabrosayev I.V., Rastyagayev D.V. Detection of Heart Rhythm Disturbances Using Analytical Spectra. *Journal of Radio Electronics*. 2015;12:16. (in Russ.) EDN:VHTMVJ
7. Friesen G.M., Jannett T.C., Jadallah M.A., Yates S.L., Quint S.R., Nagle H.T. A comparison of the noise sensitivity of nine QRS-detection algorithms. *IEEE Transactions on Biomedical Engineering*. 1990;37(1):85–98. DOI:10.1109/10.43620
8. Akopyan B. Development of the automated cardiac rhythm disorders detection and classification algorithm. *Bulletin of the UNESCO Department "Distance education in engineering" of the SUAI*. 2022;7:28–31. EDN:QSKIML
9. Kohler B.-U., Hennig C., Orglmeister R. The principles of software QRS detection. *IEEE Engineering in Medicine and Biology Magazine*. 2002;21(1):42–57. DOI:10.1109/51.993193.
10. Zong W., Moody G.B., Jiang D. A robust open-source algorithm to detect onset and duration of QRS-complexes. *Proceedings of the Conference on Computers in Cardiology, 21–24 September 2003, Thessaloniki, Greece*. IEEE; 2003. p.737–740. DOI:10.1109/CIC.2003.1291261
11. Zharinov O., Zharinov I. Correlation-extreme method for detection tasks solution and time points estimation of QRS-complexes in an electrocardiogram. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2011; 5(75):85–90. (in Russ.) EDN:OCBFFH
12. Bozhenko V.V., Chernysh N.Yu., Tatarnikova T.M. Data Mining in the Diagnosis of Anemia by Clinical Indicators. *Journal of Instrument Engineering.* 2024;67(4):321–329. DOI:10.17586/0021-3454-2024-67-4-321-329. (in Russ.) EDN:AUAHNY
13. Raskopina A.S., Bozhenko V.V., Tatarnikova T.M. Using Deep Learning in Pneumonia Diagnosis from X-Rays Patterns. *Journal of Instrument Engineering*. 2024;67(4):315–320. DOI:10.17586/0021-3454-2024-67-4-321-329. (in Russ.) EDN:UPSNQQ

Статья поступила в редакцию 13.10.2024; одобрена после рецензирования 13.11.2024; принята к публикации 20.11.2024.
The article was submitted 13.10.2024; approved after reviewing 13.11.2024; accepted for publication 20.11.2024.

Информация об авторе:

АКОПЯН Белла Кареновна
старший преподаватель кафедры прикладной информатики Санкт-Петербургского государственного университета аэрокосмического приборостроения
https://orcid.org/0000-0001-5298-9015

Автор сообщает об отсутствии конфликтов интересов.
The author declares no conflicts of interests.

Научная статья
УДК 004.056.53
https://doi.org/10.31854/1813-324X-2024-10-6-55-67

Мягкая биометрия для аутентификации и определения рук на основе использования клавиатуры

Юсеф Мохаммед Абд Аллх Альотум ✉, yousefot49@gmail.com
Андрей Владимирович Красов, krasov.av@sut.ru

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

Аннотация

Актуальность. *В настоящее время технологические системы, искусственный интеллект, общедоступность Интернета и проникновение злоумышленников в системы банков, учреждений и социальных сетей стали изучаемой наукой и доступны для всех групп и возрастов. Одна из основных задач – обеспечение системы защиты конфиденциальной информации от хакеров, а также простого доступа к аутентификации и идентификации пользователей. На первый план вышли биометрические системы, в том числе динамика движения мыши и динамика нажатия клавиш, которые выявляют стиль набора и движения мыши у каждого человека. Мягкая биометрия – интересный и недорогой биометрический метод, не требующий дополнительного оборудования. Система идентифицирует человека на основе ввода им информации в специальной графе. Динамика идентификации руки попадает в категорию поведенческой мягкой биометрии, то есть паттерны пользователя отражают индивидуальную программу действий, которой он следует при использовании сайта.*

Цель *настоящей работы – разработка системы усиленной аутентификации.*

Методы исследования. *При выполнении работы использовались методы анализа и синтеза, теории алгоритмов, законы кинематики, нейронные сети, динамика нажатия клавиш и мягкая биометрия.*

Результаты*. Описан метод извлечения динамических характеристик нажатия клавиш. Создана нейронная сеть и определено пороговое значение для выявления типа печатающей руки.*

Научная новизна. *В отличие от известных способов аутентификации, предлагаемый метод используется для определения печатающей руки на клавиатуре через нейронную сеть с помощью законов кинематики, мягкой биометрии и извлечения динамики нажатия клавиш с целью определения ценности и точности определения типа печатающей руки.*

Значимость. *Предложенное решение позволяет повысить безопасность аутентификации пользователей, увеличить скорость внедрения и снизить стоимость нового способа верификации. Результаты, полученные в работе, являются положительными и могут быть использованы в ближайшем будущем. В свою очередь, мягкие биометрические измерения зависят от поведенческих паттернов человека, что усложняет фальсификацию пользователя. Имитировать поведение при наборе текста сложно, поскольку оно является баллистическим (полуавтономным), что делает поведенческую информацию ценной, в качестве мягкого и чувствительного биометрического метода.*

Ключевые слова: *мягкая биометрия, идентификации руки, биометрическая аутентификация, закон расстояния и скорости Ньютона*

Ссылка для цитирования: Альотум Ю.М.А.А., Красов А.В. Мягкая биометрия для аутентификации и определения рук на основе использования клавиатуры // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 55–67. DOI:10.31854/1813-324X-2024-10-6-55-67. EDN:BGOBWS

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-55-67

Soft Biometrics for Authentication and Identification Hand Based on the Use of the Keyboard

Yousef M.A.A. Alotoum ✉, yousefot49@gmail.com
Andrey V. Krasov, krasov.av@sut.ru

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

Annotation

Relevance. *Nowadays, technological systems, artificial intelligence, the general availability of the Internet and penetration into the systems of banks, institutions and social networks have become a studied science and are accessible to all groups and ages. One of the main tasks was to provide a system for protecting confidential information from hackers, as well as easy access to authentication and identification of users. Biometric systems came to the fore, including mouse movement dynamics and keystroke dynamics, which reveal the typing style and mouse movement of each person. Soft biometrics is an interesting and inexpensive biometric method that does not require additional equipment. The system identifies a person based on the input information they enter in a special column. Hand identification dynamics falls into the category of behavioral soft biometrics, that is, the user's patterns reflect the individual program of actions that he follows when using the site.*

The goal of this article *the purpose of this work is to improve the security level by creating a function that will strengthen the authentication system and improve the iron gate*

Методы исследования. *In carrying out the work, methods of analysis and synthesis, theories of algorithms, laws of kinematics, neural networks, keystroke dynamics and soft biometrics were used.*

Results*. A method for extracting dynamic characteristics of keystrokes is described. A neural network is created and a threshold value is determined for identifying the type of typing hand.*

Scientific novelty. *Unlike known authentication methods, the proposed method is used to determine the typing hand on the keyboard through a neural network using the laws of kinematics, soft biometrics and extracting the dynamics of keystrokes in order to determine the value and accuracy of determining the type of typing hand.*

Significance. *The proposed solution allows to increase the security of user authentication, increase the speed of implementation and reduce the cost. The results obtained in the work are positive and can be used in the near future. In turn, soft biometric measurements depend on human behavioral patterns, which complicates user falsification. It is difficult to imitate typing behavior, since it is ballistic (semi-autonomous), which makes behavioral information valuable as a soft and sensitive biometric method.*

Keywords: *soft biometrics, hand identification, biometric authentication, Newton's law of distance and speed*

For citation: Alotoum Y.M.A.A., Krasov A.V. Soft Biometrics for Authentication and Identification Hand Based on the Use of the Keyboard. *Proceedings of Telecommunication Universities.* 2024;10(6):55–67. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-55-67.EDN:BGOBWS

Введение

В настоящее время технологические системы, искусственный интеллект, легкая доступность Интернета и проникновение злоумышленников в системы банков, учреждений и социальных сетей стали изучаемой наукой и доступны для всех групп и возрастов. Чтобы повысить безопасность персональных данных, исследователи включают метод, используемый для подтверждения личности человека – физические или поведенческие характеристики последнего, именуемый биометрией, в свои системы безопасности [1–3]. Мягкая биометрия представляет собой некоторую информацию о человеке, которая постоянно дополняется и изменяется вместе с поведением человека, например, измерения возраста, роста, веса или бихевиористические особенности. Жесткая биометрия анализирует постоянные биометрические данные и идентифицирует пользователя по отпечаткам пальцев,

изображению лица и т. д. Жесткая биометрия влечет за собой дополнительные затраты на оборудование и, следовательно, снижает готовность пользователей применять соответствующий механизм проверки. Чтобы решить эту проблему, исследователи предложили использовать определение динамики нажатия клавиш – это поведенческая биометрическая модальность для защиты от несанкционированного доступа к учетной записи. Технология динамики нажатия клавиш применяется в целях аутентификации людей, подчеркивая, что каждый пользователь печатает на клавиатуре характерным образом [4, 5]. Аутентификация пользователей по клавиатурному почерку основывается на анализе особенностей набора текста каждого пользователя. Скорость набора, длительность удержания клавиш, интервалы между нажатиями и другие параметры могут быть частью этих функций. Собирая и анализируя данные, система создает для каждого пользователя индивидуальный «клавиатурный отпечаток», который можно использовать для аутентификации. Клавиатурный почерк трудно подделать, поскольку он зависит от физиологических и поведенческих характеристик пользователя [6]. Поскольку клавиатура является основным средством ввода информации в компьютер, аутентификация на основе динамики нажатия клавиш не требует дополнительных затрат на оборудование.

Необходимо было создать систему, способную самодополняться за счет изменения динамики нажатия клавиш биометрической клавиатуры для определения печатающей руки (одна рука или две), поскольку у каждого человека свой стиль работы на клавиатуре, отличающийся от стиля другого человека с точки зрения скорости, местоположения, движения и интересов человека внутри системы. Эта система позволит существенно сократить количество подобных кибератак [7, 8].

Биометрический метод

Биометрия, в отличие от других методов проверки, опирается на характеристики, присущие каждому человеку [6, 9, 10]. Существует две категории биометрии: физиологическая и поведенческая. В первом случае для подтверждения идентификации используются физические черты тела человека, такие как лицо, отпечатки пальцев, вены или радужная оболочка глаз. Мы рассматриваем приобретенное поведение в поведенческой биометрии. В этом случае мы проверяем человека, наблюдая, как он научился выполнять определенное действие уникальным, стандартизированным способом [11–14]. Современные исследователи, проанализировав данные Джейна Эт Ала (1998), сделали вывод, что биометрическая система должна соответствовать семи различным критериям: универсальность, уникальность, постоянство, собираемость, производительность, принятие и обход. Уникальность и постоянство – два элемента, которые имеют решающее значение для оценки эффективности. В то время как постоянство связано с необходимостью иметь возможность идентифицировать человека в течение более длительного периода времени, уникальность относится к необходимости различать двух разных людей [1, 15–17].

Основные преимущества биометрии: нет необходимости запоминать пароли или использовать другие элементы / токены, предоставляющие доступ к ресурсам; повышается безопасность; биометрические данные могут использоваться для защиты от некоторых мошеннических атак, например, фишинга [1, 2, 17]. Биометрическая система реализуется посредством аутентификации и идентификации.

Аутентификация – это процедура проверки подлинности пользователя; может быть статической (система проверяет пользователя только один раз в начале сеанса) и непрерывной или активной (система контролирует пользователя на протяжении всего сеанса, чтобы обнаружить любые изменения личности во время этого сеанса [2, 5, 17, 18]).

Идентификация – это процедура установления личности пользователя.

Мягкая биометрия

Биометрические характеристики, которых недостаточно для аутентификации пользователя, такие как рост, пол, кожа, глаза, цвет волос, и которые основаны на различиях черт людей (уникальное представление личности), и эти характеристики доступны всем [15, 19–21]. Мягкая биометрия позволяет уточнить поиск реального пользователя в базе данных, что приводит к сокращению вычислительного времени. Например, если результат «захвата» биометрических данных определяет, что посетитель сайта – мужчина, согласно единице мягкой биометрии, стандартная система биометрической аутентификации может ограничить поле поиска до пользователя-мужчины без учета женщин. Мягкая биометрия для аутентификации позволяет определять эмоциональное состояние (может быть обнаружено на 84 %), пол (на 90 %), одной или двумя руками набран текст на клавиатуре (на 80 %). Наиболее многообещающие результаты аутентификации по мягкой биометрии основаны на классификации уверенности, нерешительности, нервозности, расслабления, печали и усталости с точностью от 77 до 88 %, определении, левша или правша пользователь, и определении возрастной группы. Большинство методов аутентификации пользователя ориентированы на то, когда пользователь инициирует сеанс только во время входа в систему, но также важно его аутентифицировать во

время сеанса, поэтому появляется термин «непрерывная аутентификация» [20, 22–24].

Одна из задач данного исследования – извлечь как можно больше биометрических признаков. В целях повышения точности аутентификации необходимо создать алгоритм, основанный на определении руки, которой пользователь набирает текст на клавиатуре (правой, левой, обеими руками).

Клавиатура qwerty содержит 56 клавиш, которыми можно ввести пароль. В проекте клавиатура разделена на восемь частей, как показано на рисунке 1, и состоит из следующих букв, цифр и символов:

1) первая левая часть – **~**, **1**, **2**, **3**, **4**, **5**, а также другие символы при нажатии клавиши Shift (**~**, **!**, **@**, **$**, **%**);

2) первая правая часть – **6**, **7**, **8**, **9**, **0**, **-**, **=**, а также другие символы при нажатии клавиши Shift (**^**, **&**, *, **(**, **)**, **_** ,**+**);

3) вторая левая часть – **Tab**, **q**, **w**, **e**, **r**, **t**;

4) вторая правая часть – **y**, **u**, **I**, **o**, **p**, **[**,**]**, \, а также другие символы при нажатии клавиши Shift (**{**, **}**, |);

5) третья левая часть – **Capslook**, **a**, **s**, **d**, **f**, **g**);

6) третья правая часть – **h**, **j**, **k**, **l**, **;**, **'**, а также другие символы при нажатии клавиши Shift (**:**, **"**);

7) четвертая левая часть – **left_shift**, **z**, **x**, **c**, **v**;

8) четвертая правая часть – **b**, **n**, **m**, **„**, **.** , **/**, **right_shift**, а также другие символы при нажатии клавиши Shift (**<**, **>**, **?**).

Независимо от того, какая задействована рука (правая или левая), скорость набора одной рукой заметно ниже, чем при использовании обеих рук, так как в последнем случае расстояние, которое человек проходит при переключении с одной буквы на другую, короче, чем при наборе одной рукой. Исследователи А. Перейра, Д.Л. Ли, Х. Садишкумар, Ч. Ларош, Д. Оделл и Д. Рempел опубликовали результаты исследования, проведенного с целью изучения влияния расстояния между клавишами на скорость набора текста, количество допускаемых ошибок, удобство использования, активность мышц предплечья и положение запястья. Исследование было сосредоточено на конструкции традиционной механической клавиатуры, а не на экранной, создаваемой программным обеспечением [16, 25, 26]).

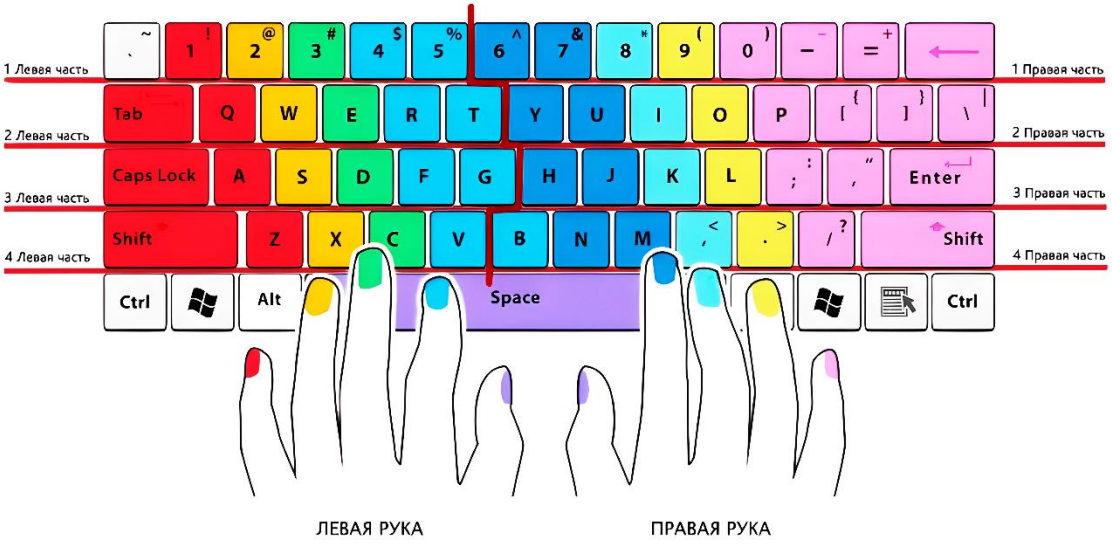


Рис. 1. Части клавиатуры

Fig. 1. Keyboard Parts

На основании исследований А. Перейра и других ученых в 1970-х гг. в работе [18] указано, что на обычное расстояние между клавишами на клавиатуре больше влияет отраслевая практика, чем вопросы эргономики (скорость набора текста, биомеханика, частота ошибок и удобство использования). Исследователи отмечают, что Международная организация по стандартизации (ISO), Американский национальный институт стандартов и Общество человеческого фактора и эргономики (ANSI/HFES) рекомендуют, чтобы горизонтальное и вертикальное межцентровые расстояния (при взгляде на клавиатуру сверху) составляли 19 мм + / –1 мм, хотя не все, но большинство конструкций клавиатуры соответствуют этим стандартам.

Взаимосвязь производительности и расстояния между клавишами была рассмотрена в исследованиях А. Перейра и др. исследователей. Так, в Японии исследователи пришли к выводу, что у людей с маленькими пальцами не наблюдается снижение производительности при различном расстоянии между клавишами (диапазон – от 15 до 19,7 мм); в то же время производительность действительно была снижена у людей с большими пальцами, если

расстояние между клавишами составило 16 мм или меньше (исследователи предостерегают от приложения выявленных результатов к населению США или других стран из-за различий в антропометрии рук) [18, 25]. В работе [27] показано, что увеличено время ввода и частота ошибок с использованием цифровых клавиатур при увеличении расстояния с 19 до 21 мм. Обзор литературы 1972 г., в котором рассматривались параметры конструкции клавиатуры того времени, дал основание полагать, что оптимальное расстояние между центрами клавиш составляет 18,1 мм [22]. В работах [28, 29] было выявлено, что при совместном рассмотрении скорости набора текста, частоты ошибок и предпочтений пользователя интервал в 19 мм был лучшим (по сравнению с другими: 14,3; 16,6; 21,4).

Ни в одном из вышеупомянутых исследований не рассматривалось *ключевое* влияние расстояния на биомеханические или физиологические показатели. Понимая, что люди с пальцами меньшей длины скорее всего лучше адаптируются к сокращению расстояния между клавишами, А. Перейра сосредоточил внимание на людях с более длинными пальцами, и в статье [19] автор в первую очередь исследует горизонтальное расстояние между клавишами.

Алгоритм идентификации руки

Каждой цифре, букве или символу на клавиатуре присвоено значение в шестнадцатеричной системе, и соответственно было создано 8 матриц на основе разделения клавиатуры на части (см. рисунок 1).

Американский стандартный код обмена информацией (ASCII, *аббр. от англ.* American Standard Code for Information Interchange) представляет собой стандарт кодирования символов для электронной связи. Коды ASCII представляют текст в компьютерах, телекоммуникационном оборудовании и других устройствах. В связи с техническими ограничениями компьютерных систем на момент его изобретения ASCII имеет всего 128 кодовых точек, из которых только 95 являются печатными символами, что серьезно ограничивает его возможности. Современные компьютерные системы используют Unicode, который имеет миллионы кодовых точек, но первые 128 из них совпадают с набором ASCII [2, 11, 26, 30].

ASCII частично был разработан на основе телеграфного кода. Его первое коммерческое использование было в Teletype Model 33 и Teletype Model 35 в качестве семибитного кода телетайпа, продвигаемого службами передачи данных Bell. Работа над стандартом ASCII началась в мае 1961 г. с первого заседания подкомитета X3.2 Американской ассоциации стандартов (ныне Американский национальный институт стандартов – *сокр*. ANSI). Первое издание стандарта было опубликовано в 1963 г., в 1967 г. – было серьезно переработано, последнее обновление произошло в 1986 г. По сравнению с более ранними телеграфными кодами, предлагаемые коды Bellu Piece и ASCII были упорядочены для более удобной сортировки (т. е. расстановки в алфавитном порядке) списков и дополнительных функций для устройств, отличных от телетайпов [22, 26–28].

Первоначально основанный на (современном) английском алфавите ASCII кодирует 128 указанных символов в семибитные целые числа, как показано на рисунке 2.

95 закодированных символов можно распечатать: к ним относятся цифры от **0** до **9**, строчные буквы от **a** до **z**, прописные буквы от **A** до **Z** и символы пунктуации. Кроме того, исходная спецификация ASCII включала 33 непечатаемых управляющих кода, созданных в моделях телетайпов; большинство из них уже устарели, хотя некоторые из них все еще широко используются, например, возврат каретки, перевод строки и коды табуляции [8, 18].

Код ASCII используется для расчета значения характеристики динамики нажатия клавиш. Например, строчная буква **e** будет представлена в кодировке ASCII как двоичное число 1101001 = шестнадцатеричное 69 (**e** – девятая буква) = десятичное 105.

Значения символьного кода массива:
– первая левая часть (192, 49, 50, 51, 52, 53);
– первая правая часть (54, 55, 56, 57, 48, 189, 187);
– вторая левая часть (9, 81, 87, 69, 82, 84);
– вторая правая часть (89, 85, 73, 79, 80, 219, 221, 220);
– третья левая часть (20, 65, 83, 68, 70, 71);
– третья правая часть (72, 74, 75, 76, 186, 222);
– четвертая левая часть (16, 90, 88, 67, 86);
– четвертая правая часть (66, 78, 77, 188, 190, 191, 16).

Начальное значение создается для каждого положения кнопки на клавиатуре:
– первая левая часть (1, 2, 3, 4, 5, 6);
– первая правая часть (7, 8, 9, 10, 11, 12, 13);
– вторая левая часть (1, 2, 3, 4, 5, 6);
– вторая правая часть (7, 8, 9, 10, 11, 12, 13, 14);
– третья левая часть (1, 2, 3, 4, 5, 6);
– третья правая часть (7, 8, 9, 10, 11, 12);
– четвертая левая часть (1, 2, 3, 4, 5);
– четвертая правая часть (6, 7, 8, 9, 10, 11, 12).

Однако при случайном и беспорядочном нажатии кнопок клавиатуры (например, кликнуть клавишу «**Q**», затем – «**K**», а после – «**Z**») сложно найти общее расстояние между нажатыми пользователем клавишами. Основываясь на законах кинетической физики, расстояние определяется, как сумма полного движения тела, независимо от направления движения, совершаемого этим телом. Выходит, расстояние является стандартной величиной, а также его можно определить, как длину – путь между

начальной и конечной точками. Т. е. итоговое расстояние можно определить, измерив дистанцию, соединяющую каждую клавишу, использованную при движении рук / руки от начальной к конечной точки.

Key	Code
backspace	8
tab	9
enter	13
shift	16
ctrl	17
alt	18
pause/break	19
caps lock	20
escape	27
page up	33
page down	34
end	35
home	36
left arrow	37
up arrow	38
right arrow	39
down arrow	40
insert	45
delete	46
0	48
1	49
2	50
3	51
4	52
5	53
6	54
7	55
8	56
9	57
a	65
b	66
c	67
d	68

Key	Code
e	69
f	70
g	71
h	72
i	73
j	74
k	75
l	76
m	77
n	78
o	79
p	80
q	81
r	82
s	83
t	84
u	85
v	86
w	87
x	88
y	89
z	90
left window key	91
right window key	92
select key	93
numpad 0	96
numpad 1	97
numpad 2	98
numpad 3	99
numpad 4	100
numpad 5	101
numpad 6	102
numpad 7	103

Key	Code
numpad 8	104
numpad 9	105
multiply	106
add	107
subtract	109
decimal point	110
divide	111
f1	112
f2	113
f3	114
f4	115
f5	116
f6	117
f7	118
f8	119
f9	120
f10	121
f11	122
f12	123
num lock	144
scroll lock	145
semi-colon	186
equal sign	187
comma	188
dash	189
period	190
forward slash	191
grave accent	192
open bracket	219
back slash	220
close braket	221
single quote	222

Рис. 2. ASCII коды символов и клавиш

Fig. 2. ASCII Character Codes and Key Codes

Стандартное расстояние между каждой клавишами клавиатуры составляет 19 мм, в связи с чем расстояние, которое проходят руки при вводе пароля, должно рассчитываться от начала до конца слова. Чтобы изначально рассчитать общее пройденное расстояние, необходимо определить полную скорость кликов на клавиши, как показано на рисунке 3.

Скорость в физике делится на стандартную и векторную. *Стандартная* выражает время, необходимое объекту для прохождения определенного расстояния без указания направления. Это стандартная физическая величина, которая выражается только в количестве. Она бывает двух типов: средняя и мгновенная стандартная скорость. Первая определяется путем деления расстояния на общее время. Вторая характеризует движение в определенный момент времени. *Векторная скорость* выражает скорость, необходимую объекту для перемещения на определенное расстояние и в определенном направлении.

Рис. 3. Расстояние, скорость и время на клавиатуре

Fig. 3. Distance, Speed and Time on Keyboard

В этом исследовании учитывалась средняя стандартная скорость, поскольку объекты перемещаются для определения расстояния и времени без определения направления. Пользователь кладет руки на клавиатуру и начинает нажимать одну клавишу за другой (зависит от длины пароля). Направление нажатых пользователем клавиш не указывается, поскольку клавиатура разделена не на одну строку, а на несколько строк и столбцов, а также из-за средней стандартной скорости, рассчитываемой путем деления пройденного за путь расстояния на общее время, необходимое для прохождения этого пути, расстояние и время.

Стандартную скорость для обеих рук можно выявить, определив расстояние между кнопками при их нажатии и отпускании, чтобы расстояние при использовании обеих рук составляло примерно 19 мм, разделенных на временну́ю метку (*от англ.* TimeStamp – числовое представление текущего времени; уникальный идентификатор, который отмечает точный момент, когда произошло событие или было выполнено определенное действие; рассчитывается с точностью до миллисекунд или быстрее, в зависимости от времени использования устройства, с помощью метода «Date.new» «getTime»). Временна́я метка используется в различных приложениях, таких как ведение журнала, отладка или измерение временны́х интервалов.

Получить временну́ю метку можно методом Date().getTime(), который возвращает примитивное значение объекта Date, представляющее собой количество миллисекунд, прошедших с 1 января 1970 г., 00:00:00 UTC.

Средняя стандартная скорость печати обеими руками определяется следующим образом:

$$a = \frac{19\ \text{mm}}{b}, \quad (1)$$

где a – средняя стандартная скорость для обеих рук; b – временна́я метка.

Местоположение каждого из нажатия клавиш на клавиатуре определяется путем разделения клавиатуры на 8 частей, четыре строки и два столбца в виде матрицы:

$$rk = Key_{z.c}, \quad (2)$$

где rk – расположение клавиш; z – номер матрицы; c – расположение кнопки в матрице.

Высчитывается расположение всех букв, которые были введены в качестве пароля. Получается значение, равное абсолютной величине вычитания расстояния первой буквы от второй буквы подряд:

$$ark = |\, rk_{i.1} - rk_{i+1} \,|, \quad (3)$$

где ark – все расположение клавиш; i – расположение кнопки в матрице.

Пройденное обеими руками расстояние (рисунок 4), рассчитывается по формуле:

$$\mathrm{P}_{\text{two hand}} = \sum rt * \sum b, \quad (4)$$

где rt – расположение кнопки в матрице.

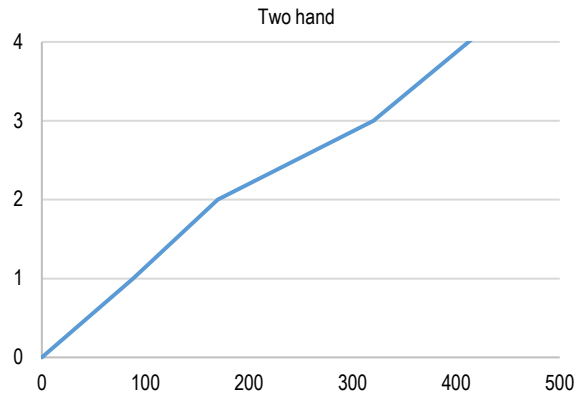


Рис. 4. Диаграмма расстояния, пройденного обеими руками

Fig. 4. Diagram of Distance Traveled by Both Hands

Средняя стандартная скорость печати одной рукой определяется следующим образом:

$$co = \frac{19\ \text{mm} * rk_{i.1}}{b}. \quad (5)$$

Расположение клавиш (расстояния между буквами соответствует 19 мм), нажатых одной рукой, которые были введены в качестве пароля, рассчитывается следующим образом:

$$rk_{\text{one hand}} = 19\ \text{mm} * rk_i\,. \quad (6)$$

Пройденное одной рукой расстояние (рисунок 5) определяется по выражению (7).

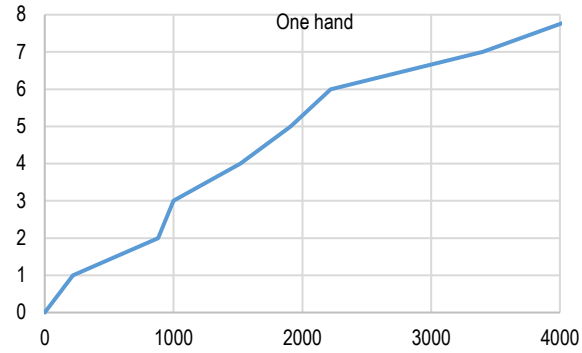


Рис. 5. Диаграмма расстояния, пройденного одной рукой

Fig. 5. Diagram of the Distance Traveled with one Hand

$$\text{P}_{\text{one hand}} = \sum co * \sum b. \quad (7)$$

Во время введения пароля одной или двумя руками, образуется сложная сеть клавиш, включающая систему параметров: точки клика, скорость, расстояние и время (рисунок 6). Чтобы узнать, как пользователь вводит пароль (одной или двумя руками), вычисляется общая скорость и расстояние от начальной до конечной точки.

Вычислить расположения букв при наборе обеими руками можно следующим образом:

$$ras = \frac{\sum rth}{k}, \quad (8)$$

где rth – расстояние, пройденное при наборе текста двумя руками; k – количество временны́х меток.

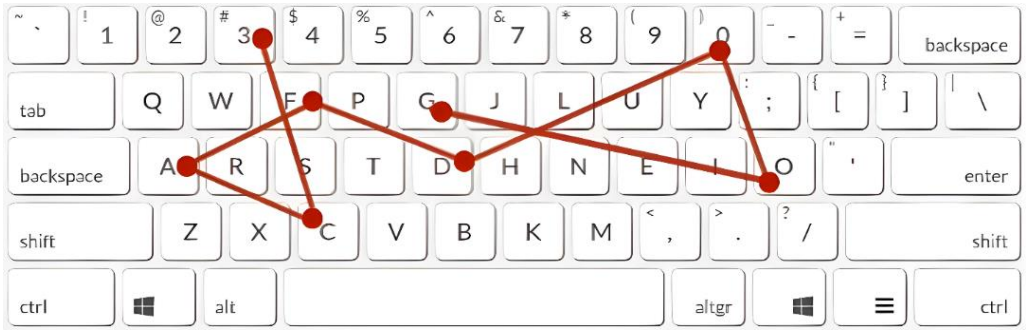


Рис. 6. Сеть клавиш клавиатуры

Fig. 6. Keyboard Button Grid

Общее расстояние, пройденное при наборе текста двумя руками (между буквами – также 19 мм), определяется как:

$$rth = \sum 19\ \text{mm} * k. \quad (9)$$

Таким образом, скорость прохождения дистанции при наборе текста двумя руками можно представить в следующем виде:

$$ste = |\, rth - ras\,|, \quad (10)$$

где ste – скорость печати двумя руками.

Расположения букв при наборе пароля одной рукой вычисляется по формуле:

$$lon = \frac{\sum ro}{k}, \quad (11)$$

где lon – расположение одной руки; ro – расстояние, пройденное при наборе текста одной рукой.

Точное значение или фиксированное расстояние, пройденного одной рукой (ФКР), можно рассчитать по выражению:

$$\text{ФКР} = k * 10. \quad (12)$$

Скорость прохождения расстояния при вводе пароля одной рукой определяется как:

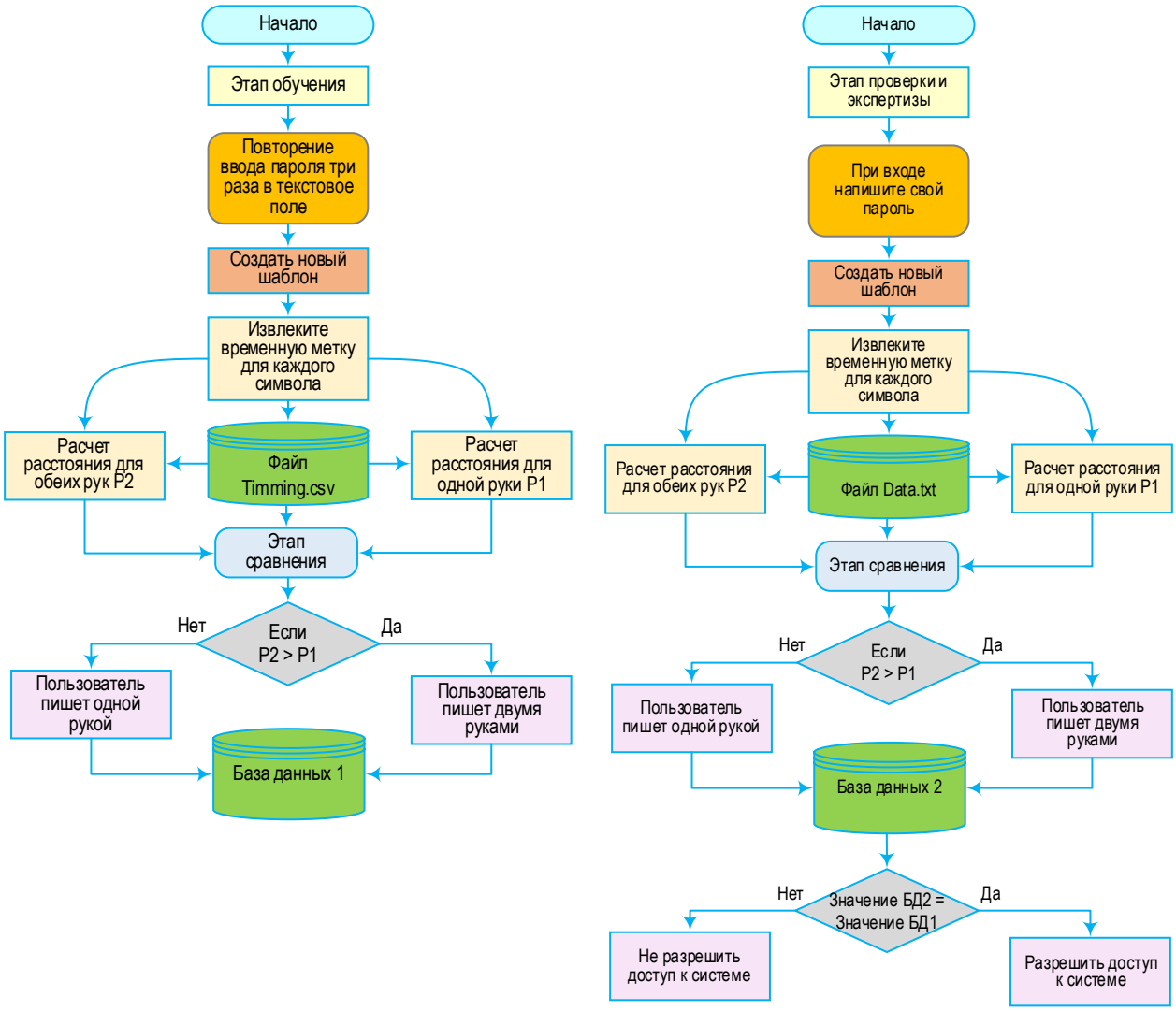
$$cpdo = | \frac{roh_{i.1}}{\sqrt{\sum b} - \text{ФКР}} |, \quad (13)$$

где *roh* – расстояние, пройденное одной рукой.

Скорость преодоления расстояния от одной клавиши до другой двумя руками выше, чем скорость преодоления этого же расстояния одной рукой. Соответственно, ориентируясь на затраченное время, можно определить одну или две руки человек использует.

Рабочая среда

Жизненный цикл системы идентификации руки при наборе текста на клавиатуре делится на два этапа: *обучение* (рисунок 7a) и *тестирование* (рисунок 7b).



a) b)

Рис. 7. Тренировка по определению почерка на этапе обучения (a) и тестирования (b)

Fig. 7. Handwriting Recognition Training at the Training Stage (a) and Testing Stage (b)

На этапе обучения пользователь трижды вводит пароль в текстовое поле и использует один и тот же метод написания одной или двумя руками во всех полях. После этого для пользователя создается специальная форма для извлечения функции нажатия клавиши и временна́я метка нажатия и отпускания клавиш. Функции нажатия клавиш и временна́я метка сохраняются в файле под названием Timing, после чего эти функции извлекаются из файла. Затем рассчитывается расстояние, пройденное одной и двумя руками, представленное сетью кнопок клавиатуры, которые были нажаты для ввода пароля в трех полях. После этого происходит этап сравнения. Если скорость прохождения дистанции двумя руками больше скорости провождения дистанции

одной рукой, то человек печатает двумя руками, а если меньше, то – одной.

На этапе тестирования пользователь вводит пароль во время входа в систему, и использует один и тот же метод написания одной или двумя руками. После этого для пользователя создается специальная форма для извлечения функции нажатия клавиш и временна́я метка нажатия и отпускания клавиш. Функции нажатия клавиш и временна́я метка сохраняются в файле под названием «Data», после чего эти функции извлекаются и рассчитывается расстояние, пройденное одной и двумя руками в сети клавиш клавиатуры, представленной кнопками, которые были нажаты для ввода пароля. В базе данных сохраняется информация о том, что расстояние, пройденное в системе клавиш двумя руками, больше, чем пройденное одной рукой. После этого происходит этап сравнения и констатация факта набора текста двумя руками или одной рукой. Далее наступает финальный этап сравнения, где, если значение, хранящееся в базе данных на этапе обучения, равно такому же значению, хранящемуся в базе данных на этапе тестирования, пользователю разрешен вход в систему, в противном случае – нет.

Заключение

В настоящее время происходит революция развития в области искусственного интеллекта и робототехники, а благодаря доступности и простоте использования Интернета область хакерства также стала преподаваемой в университетах. Таким образом, увеличивается опасность стать жертвой взлома и кражи данных.

Текущие области аутентификации и безопасности более уязвимы для взлома из-за разнообразия методов. В такие системы входят: система распознавания лиц, система отпечатков пальцев и т. д., где хакер может создать бота или программу, которая способна клонировать лицо пользователя или отпечаток пальца. Чтобы противостоять различным методам сетевого вторжения, необходимо создать систему, способную самодополняться так, чтобы координаты рабочей среды становились переменными. Для этого в исследовании был предложен способ определения стиля пользования клавиатуры одного человека.

Динамическое определение рук в настоящее время является удобной системой благодаря таким качествам, как низкая стоимость внедрения, ненавязчивое и основано исключительно на информации о том, как человек использует клавиатуру в процессе набора текста. В статье была исследована динамика набора текста на клавиатуре для аутентификации и идентификации пользователя. Кроме того, была разработана новая модель адаптивной статистики, которая позволяет корректировать пороговое значение в ответ на различия в показателях ввода пользователем пароля. Сначала были извлечены динамические характеристики нажатия клавиш и временна́я метка каждого нажатия клавиши, после чего они были отредактированы. Этот процесс повторен три раза для обучения и извлечения соответствующего порогового значения. Помимо расчета частоты ложных отклонений и приемов, для определения общего сформированного расстояния использовались кинематические уравнения.

В исследовании приняли участие 50 человек разного возраста. Были сделаны выводы, что использование законов кинематики со стилем почерка является более точным и строгим, чем другие способы определения стиля набора текста на основе динамики нажатия клавиш, поскольку оно берет среднее значение всего набора данных (рабочая зона при использовании одной руки, умноженное максимум на 10 м/с и и при использовании обоих рук), полученного в ходе обучения, и сравнивает их друг с другом, а также выдает небольшой процент ложных ошибок.

В ходе исследования было рассмотрено множество различных биометрических методов, которые хранили бы данные системы и пользователей. Вероятность несанкционированного входа в учетную запись пользователя составляет 25 %. Из приведенной статистики можно сделать вывод, что качество классических биометрических систем недостаточно надежно. Кроме того, используемые в настоящее время системы очень дороги и требуют дополнительных времени и усилий для внедрения дополнительного оснащения, например, сканеры отпечатков пальцев, глаз и лица. В пику применяемым процедурам аутентификации был разработан алгоритм, полагающийся только на клавиатуру устройства, а не на внешние системы. Разработанная система аутентификации обеспечивает повышение безопасности до 15 % за счет предотвращения несанкционированного доступа, не требует от пользователя выполнения других внешних процедур для завершения процесса входа в систему и не вызывает затруднений при использовании.

Список источников

1. Андрианов В.И., Красов А.В., Липатников В.А. Инновационное управление рисками информационной безопасности: учебное пособие. СПб.: СПбГУТ, 2012. 396 с. EDN:QSMDNH

2. Яковлев В.А., Скачкова В.В. Автоматизация выбора графического материала для систем аутентификации пользователей на основе графического пароля // Проблемы информационной безопасности. Компьютерные системы. 2015. № 1. С. 64–73. EDN:TWHDDF

3. Миняев А.А., Красов А.В., Сахаров Д.В. Метод оценки эффективности системы защиты информации территориально-распределенных информационных систем персональных данных // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2020. № 1. С. 29–33. DOI:10.46418/2079-8199_2020_1_5. EDN:ULHTJK

4. Бирих Э.В., Груздев А.С., Камалова А.О., Сахаров Д.В. Выбор инструментов динамического анализа безопасности web-приложений для задач цифровой экономики // Защита информации. Инсайд. 2024. № 1(115). С. 42–46. EDN:RLNHWK

5. Kaixin W., Hongri L., Bailing W., Shujie H., Jia S. User Authentication and Identification Model Based on Mouse Dynamics // Proceedings of the 6th International Conference on Information Engineering (ICIE '17, Dalian Liaoning, China, 17–18 August 2017). New York: Association for Computing Machinery, 2017. Article No. 18. DOI:10.1145/3078564.3078581

6. Blaganesh P., Soniya A. A Survey of Authentication Based on Mouse Behaviours // International Journal of Advanced Information Science and Technology. 2014. Vol 3. Iss. 2. PP. 42–45. DOI:10.15693/ijaist/2014.v3i2.42-45

7. Shen C., Cai Z., Guan X. Continuous Authentication for Mouse Dynamics: A Pattern-Growth Approach // Proceedings of the International Conference on Dependable Systems and Networks (DSN 2012, Boston, USA, 25–28 June 2012). IEEE, 2012. DOI:10.1109/DSN.2012.6263955

8. 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.

9. Hinbarji Z., Albatal R., Gurrin C. Dynamic User Authentication Based on Mouse Movements Curves // Proceedings of the International Conference on Multimedia Modeling (MMM 2015, Sydney, Australia, 5–7 January 2015). Lecture Notes in Computer Science. Vol. 8936. Cham: Springer, 2015. PP. 111–122. DOI:10.1007/978-3-319-14442-9_10

10. Kasprowski P., Borowska Z., Harezlak K. Biometric Identification Based on Keystroke Dynamics // Sensors. 2022. Vol. 22. Iss. 9. P. 3158. DOI:10.3390/s22093158

11. Janakiraman R., Sim T. Keystroke Dynamics in a General Setting // Proceedings of the International Conference on Biometrics (ICB 2007, Seoul, Republic of Korea, 27–29 August 2007). Lecture Notes in Computer Science. Vol. 4642. Berlin, Heidelberg: Springer, 2007. DOI:10.1007/978-3-540-74549-5_62

12. Tsimperidis I., Arampatzis A. The Keyboard Knows About You Revealing User Characteristics via Keystroke Dynamics // International Journal of Technoethics. 2020. Vol. 11. Iss. 2. DOI:10.4018/IJT.2020070103

13. Idrus S.Z.S., Cherrier E., Rosenberger C., Mondal S., Bours P. Keystroke dynamics performance enhancement with soft biometrics // Proceedings of the International Conference on Identity, Security and Behavior Analysis (ISBA 2015, Hong Kong, China, 23–25 March 2015). IEEE, 2015. DOI:10.1109/ISBA.2015.7126345

14. Gaikwad J., Kulkarni B., Phadol N., Sarukte S. User Authentication using Keystroke Dynamics // Global Research and Development Journal for Engineering. 2018. Vol. 3. Iss. 6. PP. 58–66.

15. Hassan S.I., Selim M.M., Zayed H.H. User Authentication with Adaptive Keystroke Dynamics // International Journal of Computer Science Issues. 2013. Vol. 10. Iss. 4. No 2. PP. 127–134.

16. Bours P. Continuous keystroke dynamics A different perspective towards biometric evaluation // Information Security Technical Report. 2012. Vol. 17. Iss. 1-2. PP. 36–43. DOI:10.1016/j.istr.2012.02.001

17. Mondal S., Bours P. Combining keystroke and mouse dynamics for continuous user authentication and identification // Proceedings of the International Conference on Identity, Security and Behavior Analysis (ISBA, Sendai, Japan, 29 February – 02 March 2016). IEEE, 2016. DOI:10.1109/ISBA.2016.7477228

18. Idrus S.Z.S., Cherrier E., Rosenberger C., Bours P. Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords // Computers & Security. 2014. Vol. 45. PP. 147–155. DOI:10.1016/j.cose.2014.05.008

19. Голованов А.Л. Разработка системы аутентификации по клавиатурному почерку на основе свободных текстов // В книге: Математическое и компьютерное моделирование. сборник материалов XI Международной научной конференции, посвященной памяти В.А. Романькова (Омск, Российская Федерация, 15 марта 2024). Омск: Омский государственный университет им. Ф.М. Достоевского, 2024. С. 236–237. EDN:AAOOFW

20. Сатыбалдиева М.М. Исследование систем для идентификации пользователя на основе анализа клавиатурного почерка // Научный аспект. 2024. Т. 14. № 5. С. 1897–1903. EDN:DENDWJ

21. Ямали Д.Д. Революция в аутентификации через клавиатурный почерк // Научно-исследовательский центр "Technical Innovations". 2024. № 23. С. 114–119. EDN:NLCXWH

22. Polous K.I. Comparative analysis of biometric authentication methods Общество // Молодежь. Общество. Современная наука, техника и инновации. 2021. № 20. С. 61–63. EDN:MRRBLY

23. Семенова О.С., Фадеева К.Н. Биометрическая аутентификация и её типы // III Всероссийская научно-практическая конференция с международным участием «Цифровые технологии и инновации в развитии науки и образования» (Чебоксары, Российская Федерация, 07 апреля 2023). Чебоксары: Чувашский государственный педагогический университет им. И.Я. Яковлева, 2023. С. 186–190. EDN:ILSOXW

24. Ларионов М.Ю. Перспективы развития биометрической идентификации и аутентификации личности // Инновации. Наука. Образование. 2021. № 42. С. 897–902. EDN:QEIUXB

25. Красов А.В., Альотум Ю., Ушаков И.А., Максимов В.В., Архипов А.В. Аутентификация и идентификация пользователя с использованием биометрической динамики нажатия клавиш на основе «манхэттенского и евклидовского расстояния» // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2023. № 4. С. 49–56. DOI:10.46418/2079-8199_2023_4_10. EDN:ZBXUBO

26. Yousef M.A.A.A. Biometric and behavioral authentication and soft biometrics using keystroke and mouse dynamics // XII Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2023, Санкт-Петербург, Российская Федерация, 28 февраля – 01 марта 2023). В 4 т. СПб.: СПбГУТ, 2023. С. 70–75. EDN:QTKUGV

27. Шахин Г. Биометрия во встраиваемых системах // E-Scio. 2020. № 6(45). С. 314–320. EDN:ZYRDPR

28. Ермишева Ю.Д., Омельченко Т.А. Отдельные результаты применения программного средства аутентификации по клавиатурному почерку // НБИ технологии. 2023. Т. 17. № 1. С. 11–16. DOI:10.15688/NBIT.jvolsu.2023.1.2. EDN:EXXQQO

29. Бацких А.В., Дровникова И.Г., Рогозин Е.А. К вопросу использования новой информационной технологии, связанной с дополнительной аутентификацией субъектов доступа по клавиатурному почерку, в системах защиты информации от несанкционированного доступа на объектах информатизации органов внутренних дел // Вестник Воронежского института МВД России. 2020. № 2. С. 21–33. EDN:DDVYPU

30. Пащенко Д.В., Бальзанникова Е.А. Непрерывная идентификация пользователя по клавиатурному почерку с использованием представления на основе контекста состояний // XXI век: итоги прошлого и проблемы настоящего плюс. 2020. Т. 9. № 3(51). С. 74–79. DOI:10.46548/21vek-2020-0952-0012. EDN:MAJRDT

References

1. Andrianov V.I., Krasov A.V., Lipatnikov V.A. *Innovative Information Security Risk Management*. St. Petersburg: SPbSUT Publ.; 2012. 396 p. (in Russ.) EDN:QSMDNH

2. Yakovlev V.A., Skachkova V.V. Automatic Selection of Graphical Materials for Authentication System Based on a Graphical Password. *Information Security Problems. Computer Systems*. 2015;1:64–73. (in Russ.) EDN:TWHDDF

3. Minyaev A.A., Krasov A.V., Sakharov D.V. The Efficiency Evaluation Method of Distributed ISPd Protection System. *Vestnik of St. Petersburg State University of Technology and Design*. 2020;1:29–33. (in Russ.) DOI:10.46418/2079-8199_2020_1_5. EDN:ULHTJK

4. Birikh E.V., Gruzdev A.S., Kamalova A.O., Sakharov D.V. Selection of tools for dynamic analysis of web application security for digital economy tasks. *Zaŝita informacii. Inside.* 2024;1(115):42–46. (in Russ.) EDN:RLNHWK

5. Kaixin W., Hongri L., Bailing W., Shujie H., Jia S. User Authentication and Identification Model Based on Mouse Dynamics. *Proceedings of the 6th International Conference on Information Engineering, ICIE '17, 17–18 August 2017, Dalian Liaoning, China*. New York: Association for Computing Machinery; 2017. Article No. 18. DOI:10.1145/3078564.3078581

6. Blaganesh P., Soniya A. A Survey of Authentication Based on Mouse Behaviours. *International Journal of Advanced Information Science and Technology*. 2014;3(2):42–45. DOI:10.15693/ijaist/2014.v3i2.42-45

7. Shen C., Cai Z., Guan X. Continuous Authentication for Mouse Dynamics: A Pattern-Growth Approach. *Proceedings of the International Conference on Dependable Systems and Networks, DSN 2012, 25–28 June 2012, Boston, USA*. IEEE; 2012. DOI:10.1109/DSN.2012.6263955

8. 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.

9. Hinbarji Z., Albatal R., Gurrin C. Dynamic User Authentication Based on Mouse Movements Curves. *Proceedings of the International Conference on Multimedia Modeling, MMM 2015, 5–7 January 2015, Sydney, Australia. Lecture Notes in Computer Science, vol.8936*. Cham: Springer; 2015. p.111–122. DOI:10.1007/978-3-319-14442-9_10

10. Kasprowski P., Borowska Z., Harezlak K. Biometric Identification Based on Keystroke Dynamics. *Sensors*. 2022;22(9): 3158. DOI:10.3390/s22093158

11. Janakiraman R., Sim T. Keystroke Dynamics in a General Setting. *Proceedings of the International Conference on Biometrics, ICB 2007, 27–29 August 2007, Seoul, Republic of Korea. Lecture Notes in Computer Science, vol. 4642*. Berlin, Heidelberg: Springer; 2007. DOI:10.1007/978-3-540-74549-5_62

12. Tsimperidis I., Arampatzis A. The Keyboard Knows About You Revealing User Characteristics via Keystroke Dynamics. *International Journal of Technoethics*. 2020;11(2). DOI:10.4018/IJT.2020070103

13. Idrus S.Z.S., Cherrier E., Rosenberger C., Mondal S., Bours P. Keystroke dynamics performance enhancement with soft biometrics. *Proceedings of the International Conference on Identity, Security and Behavior Analysis, ISBA 2015, 23–25 March 2015, Hong Kong, China*. IEEE; 2015. DOI:10.1109/ISBA.2015.7126345

14. Gaikwad J., Kulkarni B., Phadol N., Sarukte S. User Authentication using Keystroke Dynamics. *Global Research and Development Journal for Engineering*. 2018;3(6):58–66.

15. Hassan S.I., Selim M.M., Zayed H.H. User Authentication with Adaptive Keystroke Dynamics. *International Journal of Computer Science Issues*. 2013;10(4):127–134.

16. Bours P. Continuous keystroke dynamics A different perspective towards biometric evaluation. *Information Security Technical Report*. 2012;17(1-2):36–43. DOI:10.1016/j.istr.2012.02.001

17. Mondal S., Bours P. Combining keystroke and mouse dynamics for continuous user authentication and identification. *Proceedings of the International Conference on Identity, Security and Behavior Analysis, ISBA, 29 February – 02 March 2016, Sendai, Japan*. IEEE; 2016. DOI:10.1109/ISBA.2016.7477228

18. Idrus S.Z.S., Cherrier E., Rosenberger C., Bours P. Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords. *Computers & Security*. 2014;45:147–155. DOI:10.1016/j.cose.2014.05.008

19. Golovanov A.L. Development of an authentication system based on keyboard handwriting based on free texts. *In: Mathematical and Computer Modeling. Collection of Materials of the XI International Scientific Conference Dedicated to the Memory of V.A. Romankov, 15 March 2024, Omsk, Russian Federation*. Omsk: Dostoevsky Omsk State University Publ.; 2024. p.236–237. (in Russ.) EDN:AAOOFW

20. Satybaldieva M.M. Research of systems for user identification based on the analysis of keyboard handwriting. *Scientific Aspect*. 2024;14(5):1897–1903. (in Russ.) EDN:DENDWJ

21. Yamali D.D. Revolution in authentication through keyboard handwriting. *Research Center "Technical Innovations"*. 2024;23:114–119. (in Russ.) EDN:NLCXWH

22. Polous K.I. Comparative analysis of biometric authentication methods Society. *Youth. Society. Modern Science, Technology and Innovation*. 2021;20:61–63. EDN:MRRBLY

23. Semenova O.S., Fadeeva K.N. Biometric authentication and its types. *Proceedings of the IIIrd All-Russian Scientific and Practical Conference with International Participation on Digital Technologies and Innovations in the Development of Science and Education, 07 April 2023, Cheboksary, Russian Federation*. Cheboksary: I. Yakovlev Chuvash State Pedagogical University Publ.; 2023. p.186–190. (in Russ.) EDN:ILSOXW

24. Larionov M.Yu. Prospects for the development of biometric identification and authentication of personality. *Innovations. Science. Education*. 2021;42:897–902. (in Russ.) EDN:QEIUXB

25. Krasov A.V., Alyotum Yu., Ushakov I.A., Maksimov V.V., Arkhipov A.V. User authentication and identification using biometric keystroke dynamics based on the "Manhattan and Euclidean distance". *Vestnik of St. Petersburg State University of Technology and Design. Series 1: Natural and Technical Sciences*. 2023;4:49–56. (in Russ.) DOI:10.46418/2079-8199_2023_4_10. EDN:ZBXUBO

26. Yousef M.A.A.A. Biometric and behavioral authentication and soft biometrics using keystroke and mouse dynamics // Proceedings of the XIIth International Conference on Infotelecommunications in Science and Education, 28 February – 01 March 2022, St. Petersburg, *Russian Federation*. St. Petersburg: The Bonch-Bruevich Saint-Petersburg State University of Telecommunications Publ.; 2023. p. 70–75. (in Russ.) EDN:QTKUGV

27. Shahin G. Biometrics in embedded systems. *E-Scio*. 2020;6(45):314–320. (in Russ.) EDN:ZYRDPR

28. Ermisheva Yu.D., Omelchenko T.A. Separate results of the application of the software authentication tool by keystroke dynamics. *NBI Technologies*. 2023;17(1):11–16. (in Russ.) DOI:10.15688/NBIT.jvolsu.2023.1.2. EDN:EXXQQO

29. Batskikh A.V., Drovnikova I.G., Rogozin E.A. On the issue of using a new information technology related to additional authentication of access subjects using keyboard handwriting in information protection systems against unauthorized access at information objects of internal affairs bodies. *The Bulletin of Voronezh Institute of the Ministry of Internal Affairs of Russia*. 2020;2:21–33. (in Russ.) EDN:DDVYPU

30. Pashchenko D.V., Balzannikova E.A. Continuous keystroke dynamics user identification using state context representation. *XXI Century: Resumes of the Past and Challenges of the Present plus*. 2020;9(3(51)):74–79. (in Russ.) DOI:10.46548/21vek-2020-0952-0012. EDN:MAJRDT

Статья поступила в редакцию 04.06.2024; одобрена после рецензирования 23.09.2024; принята к публикации 07.10.2024.
The article was submitted 04.06.2024; approved after reviewing 23.09.2024; accepted for publication 07.10.2024.

Информация об авторах:

АЛЬОТУМ Юсеф Мохаммед Абд Аллх
аспирант кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0009-0000-8684-7664

КРАСОВ Андрей Владимирович
кандидат технических наук, доцент, заведующий кафедрой защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0002-9076-6055

Авторы сообщают об отсутствии конфликтов интересов.
The authors declare no conflicts of interests.

Научная статья
УДК 519.688
https://doi.org/10.31854/1813-324X-2024-10-6-68-78

Алгоритм синтеза групп кодов в RFID-системе множественного доступа

Наталья Аркадьевна Верзун [1] ✉, verzun.n@unecon.ru
Алексей Михайлович Колбанёв [2], kolbanev@gmail.com
Михаил Олегович Колбанёв [1], mokolbanev@mail.ru

[1]Санкт-Петербургский государственный экономический университет,
Санкт-Петербург, 191023, Российская Федерация
[2]АО «ЭР-Телеком Холдинг»,
Москва, 115035, Российская Федерация

Аннотация

Актуальность. *Одной из проблем, которую необходимо решать при создании RFID-систем, является множественный доступ ридера к группе меток, расположенных в ограниченном пространстве, поскольку считывающий сигнал вызывает одновременный отклик многих меток, что приводит к коллизиям (конфликтам) ответных сигналов. Эта проблема не решена применительно к пассивным меткам без чипа, построенным на технологиях поверхностных акустических волн (ПАВ), код которых закладывается при изготовлении и не может быть изменен в процессе эксплуатации.*

Цель *проведенного исследования заключается в разработке алгоритмов, позволяющих синтезировать такие группы кодов, которые обеспечивали бы управляемый уровень попарной корреляции ответных сигналов меток и за счет этого обеспечивали бы заданную точность идентификации меток. В основе предложенных алгоритмов лежат процедуры конкатенации кодов и индуктивного построения групп кодов с заданными емкостью и уровнем корреляции. Для алгоритма формирования группы кодов с требуемым значением коэффициента корреляции и алгоритма объединения групп кодов в полные и максимальные группы доказаны свойства, подтверждающие возможность использования их для формулирования заданий на изготовление групп меток на поверхностных акустических волнах, которые соответствовали бы количеству объектов, требующих идентификации, и точности их идентификации с учетом количества меток в группе, условий распространения радиосигналов в зоне работы ридера, количества повторных считываний кодов меток, а также алгоритмов совместной обработки данных, полученных при всех считываниях. Для достижения цели исследования используются* ***методы*** *теории кодирования, корреляционного анализа.*

Результат. *Разработанный алгоритм представляет собой инструмент создания современных систем кодирования для меток на ПАВ.*

Научная новизна. *Известные алгоритмы множественного доступа в RFID-системах предложены в стандартах GEN1 и GEN2 EPC Global, и предполагают наличие у метки чипа и блока питания, что позволяет реализовывать протоколы воздействия на метку ридером при помощи специальных команд. Предлагаемый алгоритм множественного доступа применим для пассивных меток на ПАВ, в том числе, передвигающихся на высокой скорости и / или расположенных в агрессивных средах, так как метки не используют кремниевую технологию по сравнению с активными RFID-метками.*

Практическая значимость. *Использование предложенного комплекса алгоритмов позволит повысить эффективность систем маркировки за счет сокращения времени идентификации объектов, находящихся в замкнутом пространстве.*

Ключевые слова: *технология радиочастотной идентификации, маркировка объектов, группа кодов, корреляция кодов, алгоритм синтеза группы кодов*

Ссылка для цитирования: Верзун Н.А., Колбанёв А.М., Колбанёв М.О. Алгоритм синтеза групп кодов в RFID-системе множественного доступа // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 68–78. DOI:10.31854/1813-324X-2024-10-6-68-78. EDN:POITEX

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-68-78

An Algorithm for Synthesizing Groups of Codes in an RFID Multiple Access System

Nataliya A. Verzun [1] ✉, verzun.n@unecon.ru
Aleksey M. Kolbanev [2], kolbanev@gmail.com
Michail O. Kolbanev [1], mokolbanev@mail.ru

[1]St. Petersburg State University of Economics,
St. Petersburg, 191023, Russian Federation
[2]ER-Telecom Holding,
Moscow, 115035, Russian Federation

Annotation

Relevance. *One of the problems that must be solved when creating RFID systems is the reader's multiple access to a group of tags located in a limited space, since the reading signal causes a one-time response of many tags, which leads to collisions (conflicts) of response signals. This problem has not been solved in relation to passive tags without a chip, based on surface acoustic wave technologies, the code of which is laid down during manufacture and cannot be changed during operation.*

The purpose of the study *is to develop algorithms that allow synthesizing such groups of codes that would provide a controlled level of pairwise correlation of the selected label signals and thereby ensure the specified accuracy of label identification. The proposed algorithms are based on the procedures of code concatenation and inductive construction of groups of codes with a given volume and correlation level. For the algorithm for forming a group of codes with the required value of the correlation coefficient and the algorithm for combining groups of codes into complete and maximum groups, properties have been proven that confirm the possibility of using them to formulate tasks for preparing groups of labels on surfactants that would correspond to the number of objects requiring identification and the accuracy of their identification and taking into account the number of labels in the group, the conditions for the propagation of radio signals in the area of operation of the reader, the number of repeated readings of the label codes, as well as algorithms for joint data processing, received with all calculations.*

The methods used. *Methods of coding theory and correlation analysis.*

Result. *The developed algorithm is a tool for creating modern coding systems for surfactant labels.*

The scientific novelty. *Well-known algorithms for multiple access in RFID systems are proposed in the GEN1 and GEN2 EPC Global standards, and assume that the tag has a chip and a power supply, which makes it possible to implement protocols for influencing the tag with a reader using special commands. The proposed multiple access algorithm is applicable for passive surfactant tags, including those moving at high speed and/or located in aggressive environments, since the tags do not use silicon technology compared to active RFID tags.*

Practical significance. *The use of the proposed set of algorithms will increase the efficiency of marking systems by reducing the identification time of objects located in a confined space.*

Keywords: *radio frequency identification technology, object marking, code group, code correlation, code group synthesis algorithm*

For citation: Verzun N.A., Kolbanev A.M., Kolbanev M.O. An Algorithm for Synthesizing Groups of Codes in an RFID Multiple Access System. *Proceedings of Telecommunication Universities.* 2024;10(6):68–78. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-68-78. EDN:POITEX

Введение

Отдельную группу технологий интернета вещей составляют технологии радиочастотной идентификации – RFID и ее подвид NFC. Этот способ бесконтактного взаимодействия с вещами существенно отличается от других, хотя сам термин «интернет вещей» появился в 1990-е гг. благодаря именно RFID-системам [1, 2], к числу главных элементов которой относятся RFID-метки и RFID-

ридеры. Первые содержат уникальный код и сопрягаются с физическими вещами, вторые способны получить радиодоступ к метке, прочитать ее код на расстоянии и сформировать сообщение о наличии или отсутствии маркированной вещи в зоне действия ридера [3].

Одной из проблем, которую приходится решать при создании RFID-системы, является проблема множественного доступа ридера к группе меток, расположенных в ограниченном пространстве. Дело в том, что если сигнал считывателя вызывает одновременный отклик многих меток, то возможны коллизии (конфликты) ответных сигналов. Такая ситуация возникает в библиотеках, когда надо прочитать метки книг, расположенных на одной полке, в магазинах, когда требуется узнать метки товаров, находящихся в корзине покупателя, на складе, в котором надо произвести инвентаризацию, и во многих других приложениях.

Очевидный способ борьбы с коллизиями такого рода можно назвать пространственным. Он заключается в уменьшении зоны действия ридера и последовательном чтении меток «вещь за вещью» при перемещении их в зону считывания. Высокая трудоемкость и большое время последовательного доступа ко всем меткам группы практически нивелируют главные достоинства RFID-идентификации.

Другой подход применим только к RFID-меткам, которые содержат чип. Стандарт ISO 15693 для меток с чипами предусматривает использование алгоритма антиколлизии – синхронный ALOHA. Согласно данному алгоритму, считыватель задает временны́е интервалы (или слоты), и передача метки может начинаться только с началом временно́го интервала. В случае, если возникла коллизия, повторная передача происходит через случайное число интервалов. Развитием синхронного ALOHA являются алгоритмы, разработанные EPC Global (*аббр. от англ.* Electronic Product Code). Они используют способность активной метки распознать команду считывателя, предназначенную для нее. При необходимости считыватель может задержать передачу данных от уже идентифицированных меток [4].

Однако особый интерес представляют методы множественного доступа к пассивным меткам без чипа, построенным на технологиях поверхностных акустических волн (ПАВ). Такие метки имеют малые стоимость и размер, могут функционировать в широком температурном диапазоне, не подвергаются воздействию радиации, поддерживают значительную дальность считывания и обладают другими достоинствами [5, 6].

Алгоритмы работы системы радиочастотной идентификации определяется требованиями предметной области к ее функционалу [7–9]. При создании RFID-системы, использующей метки на ПАВ, надо учитывать, что считываемый код каждой метки закладывается при ее изготовлении и не может быть изменен в процессе эксплуатации. Поэтому единственным способом, позволяющим избежать коллизий ответных сигналов меток, кроме последовательного доступа, представляется изготовление группы меток с такими кодами, которые имеют малую попарную корреляцию, в идеальном случае – ортогональные коды. В случае, если коды не ортогональны, но их корреляция незначительна, точность идентификации зависит от количества меток в группе, условий распространения радиосигналов в зоне работы ридера, количества повторных считываний кодов меток, а также алгоритмов совместной обработки данных, полученных при всех считываниях.

Центральной задачей при построении RFID-систем при этих условиях является формирование заданий на изготовление такой группы меток на ПАВ, которая характеризуется определенными емкостью (количеством меток) и уровнем попарной корреляции заложенных в них кодов. Таким образом, актуальной при разработке и практическом использовании информационных систем множественного доступа, основанных на технологии радиочастотной идентификации, является задача синтеза групп дискретных кодов, которые можно было бы использовать для маркировки множеств объектов [10].

Требования к группе кодов

Можно сформулировать следующие требования к группе кодов.

1. Группы кодов должны иметь достаточную *мощность M* – число уникальных кодовых комбинаций для идентификации всей совокупности маркируемых объектов, а каждый код группы должен иметь достаточную информационную емкость для отображения в нем всей необходимой для регистрации в информационной системе данных.

2. Для корректной идентификации объектов необходимо, чтобы кодовые комбинации одной группы в совокупности (то есть каждая кодовая комбинация группы с каждой другой кодовой комбинацией этой группы) были бы как можно *менее коррелированы*. Чем ниже корреляция кодовых комбинаций группы кодов, тем будет выше вероятность безошибочного распознавания меток во время считывания при одинаковом соотношении уровней полезного сигнала и помехи [11, 12].

Рассмотрим далее задачу синтеза групп дискретных слабо коррелированных кодов для маркировки идентифицируемых объектов.

Введем следующие обозначения:
возможны два значения каждого разряда кода: –1 и 1;
n – длина кодовой комбинации (длина кода);
$A = (a_1 \dots a_i \dots a_n), B = (b_1 \dots b_i \dots b_n)$ – кодовые комбинации длины n;
a_i, b_i – значения i-х ($i = \overline{1,n}$) разрядов кодовых комбинаций A и B, каждый разряд принимает одно из двух возможных значений: $a_i, b_i \in \{1, -1\}$;
$A(n,k)$ – группа кодовых комбинаций длины n, каждая из которых отличается минимум в k разрядах от остальных кодовых комбинаций группы $A(n,k)$;
$A_i = (a_i^1 a_i^2 \dots a_i^n) \in A(n,k)$, $i = 1, 2, \dots M$, – кодовая комбинация длиной n, которая принадлежит группе $A(n,k)$;
a_i^j – значение j-го разряда кодовой комбинации A_i;
M – количество кодовых комбинаций в группе $A(n,k)$.

Группу кодов $A(n,k)$ можно представить в виде матрицы: M строк и n столбцов. Обозначим ее как $\|A(n,k)\|$.

Будем предполагать, что в группе $A(n,k)$ длина кодовых комбинаций принимает значения из ряда:

$$n = 4, 8, 16, 32, 64, 128, 256 \dots, \quad (1)$$

а число разрядов, в которых различаются коды группы $A(n,k)$:

$$k = 2, 4, 8, 16, 32, 64, 128, \dots .$$

Следует отметить, что при этом во всех случаях:
– $k > 1$, поскольку при $k = 1$ поставленная задача теряет смысл, т. к. группа кодов $A(n,1)$ является совокупностью всех кодов длины n, то есть 2^n кодов;
– $k \le n$.

В случае $k = n$ число кодов в группе равно двум: прямой и инверсный.

Оценка корреляции группы кодов

Для оценки степени корреляции двух кодовых комбинаций $A = (a_1 \dots a_i \dots a_n)$ и $B = (b_1 \dots b_i \dots b_n)$ группы $A(n,k)$ предлагается использовать коэффициент корреляции Q_{AB}, который рассчитывается следующим образом:

$$Q_{AB} = \frac{a_1 b_1 + \dots + a_i b_i + \dots + a_n b_n}{n}. \quad (2)$$

Как следует из (2), Q_{AB} принимает значения в диапазоне $[-1 \dots 1]$.

Значение $Q_{AB} = 1$ свидетельствует о том, что кодовые комбинации A и B одинаковы ($a_1 = b_1, \dots, a_i = b_i \dots, a_n = b_n$) и, соответственно, полностью коррелированы.

Если $Q_{AB} = -1$, то A и B не совпадают ни по одному разряду ($a_1 \ne b_1, \dots, a_i \ne b_i \dots, a_n \ne b_n$) и корреляции между A и B нет.

В группе $A(n,k)$ кодовые комбинации отличаются как минимум k разрядами и тогда коэффициент корреляции для этой группы кодов:

$$Q_{AB}(n,k) = \frac{n - 2k}{n}. \quad (3)$$

С учетом принятых обозначений задача синтеза групп слабо коррелированных кодов для маркировки объектов сводится к формированию таких групп кодов, которые имеют низкий коэффициент корреляции $Q_{AB}(n,k)$.

Рисунок 1 иллюстрирует зависимость коэффициента корреляции от разных значений k и n, построенную по формуле (3). В частности, видно, что для n = 128 коэффициент корреляции будет:

$Q_{AB} = 0{,}984375$ при $k = 2$;
$Q_{AB} = 0{,}96875$ при $k = 4$;
$Q_{AB} = 0{,}9375$ при $k = 8$,
$Q_{AB} = 0{,}75$ при $k = 16$
и $Q_{AB} = 0{,}5$ при $k = 32$.

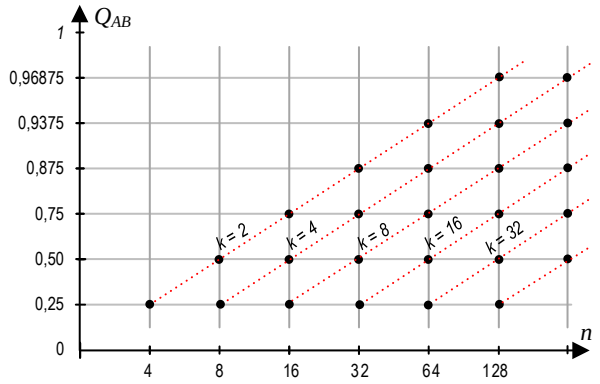


Рис. 1. Зависимость коэффициента корреляции от значений *k* и *n*

Fig. 1. The Dependence of the Correlation Coefficient on the Values of k and n

Очевидно, что, меняя параметры n и k, можно подобрать такую группу кодов, коэффициент корреляции для которой не будет превышать заданного значения. Такой «подбор» является нетривиальной задачей из-за больших значений M и n для практически значимых систем идентификации. В целях уменьшения вычислительной сложности этой задачи можно использовать две процедуры: конкатенацию двоичных кодов и индуктивное построение требуемых групп кодов. Рассмотрим их подробнее.

Конкатенация двоичных кодов

Кодовая комбинация $A_i = (a_i^1 \dots a_i^n a_i^{n+1} \dots a_i^{2n})$ – результат конкатенации кодовых комбинаций $C_v = (c_v^1 c_v^2 \dots c_v^n)$ и $C_w = (c_w^1 c_w^2 \dots c_w^n)$, если левая половина A_i равна $(a_i^1 \dots a_i^n) = (c_v^1 \dots c_v^n)$, а правая половина – $(a_i^{n+1} \dots a_i^{2n}) = (c_w^1 \dots c_w^n)$, соответственно.

Пример выполнения конкатенации представлен на рисунке 2а.

$$A_j = C_v + C_w$$
$$C_v = 010 \quad C_w = 111$$
$$A_j = 010\ 111$$

a)

$$A = C^{(1)} + C^{(2)}$$
$$C^{(1)} = \begin{vmatrix}00\\11\end{vmatrix} \quad C^{(2)} = \begin{vmatrix}01\\10\end{vmatrix}$$
$$A = \begin{vmatrix}0001\\0010\\1101\\1110\end{vmatrix}$$

b)

Рис. 2. Конкатенация двоичных кодов C_v и C_w (a) и попарная конкатенация двух групп кодов $C^{(1)}$ и $C^{(2)}$ (b)

Fig. 2. Concatenation of Binary Codes C_v and C_w (a) and Pairwise Concatenation of Two Groups of Codes $C^{(1)}$ and $C^{(2)}$ (b)

Если мы имеем две группы кодовых комбинаций, одна из которых:

$$C^{(1)} = \left\{C_1^{(1)}, \dots, C_i^{(1)}, \dots, C_M^{(1)}\right\}, \text{мощностью } M,$$

а другая:

$$C^{(2)} = \left\{C_1^{(2)}, \dots, C_i^{(2)}, \dots, C_M^{(2)}\right\}, \text{мощностью } N,$$

то результатом их попарной конкатенации будет группа из $N \cdot M$ кодов $A = \{A_1, \dots, A_j, \dots, A_{M \cdot N}\}$, где A_j – результат конкатенации каждого кода из группы $C^{(1)}$ с каждым кодом из группы $C^{(2)}$. Пример выполнения попарной конкатенации $C^{(1)}$ и $C^{(2)}$ представлен на рисунке 2b.

Для обозначения операции конкатенации будем использовать оператор «+»: $A = C^{(1)} + C^{(2)}$.

Индуктивное построение требуемых групп кодов

Целью индуктивной процедуры является формирование такой группы кодов, которая объединяла бы коды, имеющие длину $2n$, и при этом различающихся на k, $k+1$, ..., $2n$ разрядов. Построение такой группы при помощи полного перебора становится невозможным при значениях n, имеющих практическое значение. Однако при малых значениях n полный перебор возможен.

В основе построения групп с достаточно большим n и заданным коэффициентом корреляции лежит получение на первом шаге трех групп кодов с небольшими значениями n и k, а именно групп, объединяющих коды:

– длиной n, различающихся на k, $k+1$, ..., n разрядов;

– длиной n, различающихся на $k/2$, $k/2+1$, ..., n разрядов;

– длиной $2n$, различающихся на $k/2$, $k/2+1$, ..., $2n$ разрядов.

Необходимо подчеркнуть, что указанное количество разных разрядов должно выполняться для любой пары кодов, объединенных в одной группе.

Процесс завершается по достижении требуемых значений n и уровня корреляции между кодами одной группы.

Для приведенных ниже алгоритмов доказана возможность индуктивного построения групп кодов с требуемыми характеристиками.

Любая группа кодов содержит только одну группу с кодом, состоящим из одних нулей. Для удобства изложения будем называть ее базовой группой. На рисунке 3а базовой группой является группа кода $\Psi^{(1)}(4,2)$. Она состоит из 4-х групп инверсных кодов, выделенных пунктирными линиями. Сложение по модулю 2 любой пары кодов из группы $\Psi^{(1)}(4,2)$ дает код, который так же принадлежит этой группе кодов.

$$\psi^{(1)}(4,2) = \begin{vmatrix}0000 & 0\\ 1111 & 15\\ 0011 & 3\\ 1100 & 12\\ 0101 & 5\\ 1010 & 10\\ 0110 & 6\\ 1001 & 9\end{vmatrix}$$

a)

$$\psi^{(2)}(4,2) = \begin{vmatrix}0001 & 1\\ 1110 & 14\\ 0010 & 2\\ 1101 & 13\\ 0100 & 4\\ 1011 & 11\\ 1000 & 8\\ 0111 & 7\end{vmatrix}$$

b)

Рис. 3. Примеры полной и максимальной групп кодов: a) $\Psi^{(1)}(4,2)$; b) $\Psi^{(2)}(4,2)$

Fig. 3. Examples of Full and Maximum Code Groups: a) $\Psi^{(1)}(4,2)$; b) $\Psi^{(2)}(4,2)$

Остальные 8 кодов длины $n = 4$ (см. рисунок 3b) также образуют группу кодов $\Psi^{(2)}(4,2)$, отличающихся на $k = 2$, и тоже могут быть разделены на 4 группы инверсных кодов. Группа кодов $\Psi^{(2)}(4,2)$ может быть образована путем сложения по модулю 2 каждого из кодов группы $\Psi^{(1)}(4,2)$ с любым из восьми кодов, который не входит в группу $\Psi^{(1)}(4,2)$.

Введем необходимые определения.

Полную группу кодов длины n, отличающихся друг от друга как минимум в k разрядах, образует такая группа кодов, которая не может быть увеличена за счет добавления в нее еще хотя бы одного кода длины n.

Максимальная полная группа кодов – полная группа кодов с максимальной мощностью. Группы кодов $\Psi^{(1)}(4,2)$ и $\Psi^{(2)}(4,2)$, представленные на рисунке 3, являются полными и максимальными. Мощность полной и максимальной группы $\Psi(n,k)$ обозначим через $M_{n,k}^{\text{Код}}$. В примерах (см. рисунок 3) мощность $M_{4,2}^{\text{Код}} = 8$.

Полная совокупность групп кодов объединяет все группы кодов длины n, и таким образом, включает 2^n различных кодов. В примере (см. рисунок 3) объ-

единение групп кодов $\Psi^{(1)}(4,2)$ и $\Psi^{(2)}(4,2)$ образует полную совокупность групп кодов $\widehat{\Psi}(4,2)$: $\widehat{\Psi}(4,2) = \{\Psi^{(1)}(4,2), \Psi^{(2)}(4,2)\}$.

Число полных и максимальных групп кодов в полной совокупности кодов длиной n, отличающихся не менее чем на k разрядов, будем обозначать $C_{n,k}^{\text{Гр}}$. Если $n = 4$ и $k = 2$, то, как видно из рисунка 3, $C_{4,2}^{\text{Гр}} = 2$. Процесс формирования групп кодов с требуемыми свойствами включает три алгоритма:

Алгоритм 1 в качестве исходных данных использует группу кодов длиной n, отличающихся в совокупности не менее, чем k разрядами, и при помощи операции конкатенации формирует новую группу кодов длиной $2n$, которые также отличаются k разрядами. Фактически, этот алгоритм, за счет увеличения длины кодов при зафиксированном значении k, повышает число кодовых комбинаций в группе кодов M, и при этом растет их взаимная попарная корреляция. Если при достигнутом значении M коэффициент корреляции больше заданного в исходных данных, необходимо увеличить значение k и повторить выполнение алгоритма 1.

Алгоритм 2 объединяет уже сформированные группы кодов в полные и максимальные группы, которые по-прежнему отличаются в совокупности не менее, чем k разрядами. Фактически, совместная работа первого и второго алгоритмов позволяет сформировать начальную группу кодов определенной длины, отличающихся на определенное число разрядов, причем эта начальная группа является полной и максимальной.

Алгоритм 3 формирует новые полные и максимальные группы кодов, которые отличаются от начальной группы, что позволяет получить полную совокупность групп кодов. Для этого используется процедура инверсии одного или нескольких столбцов начальной группы кодов.

Алгоритм 1 формирования группы кодов с требуемым значением коэффициента корреляции

Рассмотрим процесс формирования групп кодов с требуемыми свойствами. В качестве исходных данных задаются значения свойств группы кодов: M^* – мощность группы кодов или n^* – длина группы кодов (выбирается из ряда чисел (1)), $Q_{AB}(n^*, k)$ – минимально допустимый коэффициент корреляции.

Допустим, заданы значения: n^* и $Q_{AB}(n^*, k)$, в этом случае по формуле (2) можно найти число отличающихся разрядов в кодах [13]:

$$k = \frac{n}{2}(1 - Q_{AB}(n^*, k)).$$

Итак, значения n^* и k задают требуемые свойства группы кодов, алгоритм формирования которой представлен на рисунке 4.

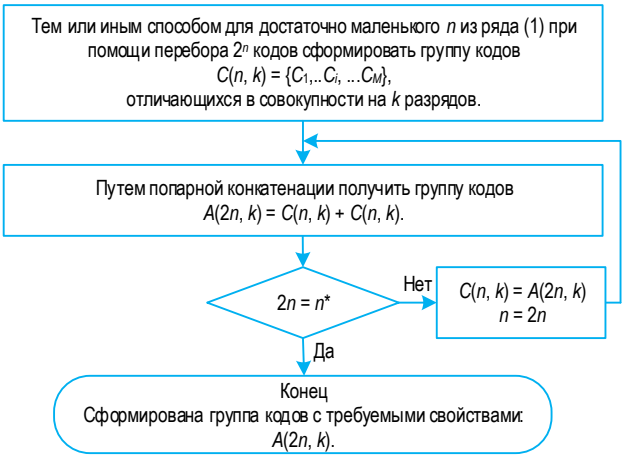


Рис. 4. Алгоритм формирования группы кодов

Fig. 4. The Algorithm for Forming a Group of Codes

Фактически, алгоритм (см. рисунок 4) за счет увеличения длины кодов и корреляции между ними увеличивает мощность множеств кодов, отличающихся определенным числом разрядов. Рассмотрим свойства группы кодов $A(2n, k)$, сформированных при помощи этого алгоритма.

Свойство 1.1. Длина кодов $A_i \in A(2n, k)$ равна $2n$, поскольку они образуются конкатенацией двух кодов групп $C(n, k)$ длиной n.

Свойство 1.2. Число кодов группы $A(2n, k)$ равно $M \cdot M = M^2$, поскольку реализуется конкатенация каждого из M кодов группы $C(n, k)$ с каждым другим кодом, включая себя самого.

Свойство 1.3. Коды группы $A(2n, k)$ отличаются в совокупности k разрядами. Чтобы убедиться в этом, выберем два разных произвольных кода:

$$A_i = (a_i^1 \dots a_i^n a_i^{n+1} \dots a_i^{2n}) \in A(2n, k),$$
$$A_j = (a_j^1 \dots a_j^n a_j^{n+1} \dots a_j^{2n}) \in A(2n, k).$$

Сравним по отдельности левые $(a_i^1 \dots a_i^n)$ и $(a_j^1 \dots a_j^n)$ и правые $(a_i^{n+1} \dots a_i^{2n})$ и $(a_j^{n+1} \dots a_j^{2n})$ части этих кодов.

Поскольку при построении кодов A_i и A_j использовалась операция попарной конкатенации кодов из множества $C(n, k)$, то:

$$(a_i^1 \dots a_i^n) \in C(n, k),$$
$$(a_j^1 \dots a_j^n) \in C(n, k),$$

и левые половины отличаются или на k разрядов, или совпадают (не отличаются ни одним из разрядов);

$$(a_i^{n+1} \dots a_i^{2n}) \in C(n, k),$$
$$(a_j^{n+1} \dots a_j^{2n}) \in C(n, k),$$

и правые половины отличаются либо на k разрядов, либо совпадают.

Если совпадают левые половины:

$$(a_i^1 \dots a_i^n) = (a_j^1 \dots a_j^n),$$

то не могут совпадать правые половины (в противном случае A_i и A_j – это один и тот же код):

$$(a_i^{n+1} \dots a_i^{2n}) \neq (a_j^{n+1} \dots a_j^{2n}).$$

Аналогично, если:

$$(a_i^{n+1} \dots a_i^{2n}) = (a_j^{n+1} \dots a_j^{2n}),$$

то

$$(a_i^1 \dots a_i^n) \neq (a_j^1 \dots a_j^n),$$

а значит, исходя из свойств кодов, объединенных в множество $C(n,k)$, коды A_i и A_j отличаются хотя бы k разрядами, и группа кодов $A(2n,k)$ обладает заявленным свойством 3.

Рассмотрим, например, коды длиной 4 бит (см. рисунок 3). Попарной конкатенацией кодов каждой из групп $\Psi^{(1)}(4,2)$ и $\Psi^{(2)}(4,2)$ могут быть получены две группы кодов $A^{(1)}(8,2)$ и $A^{(2)}(8,2)$, соответственно, по 8 · 8 = 64 кода длиной 8 бит в каждой – результат показан на рисунке 5.

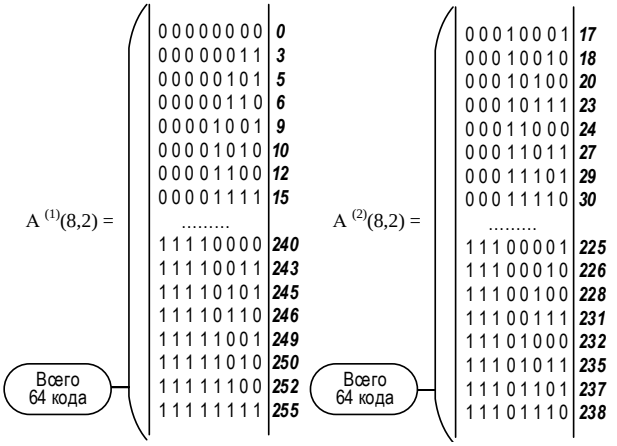


Рис. 5. Группы кодов $A^{(1)}(8,2)$ и $A^{(2)}(8,2)$

Fig. 5. Code Groups $A^{(1)}(8,2)$ and $A^{(2)}(8,2)$

На рисунке 5 показаны только первые и последние 8 кодовых комбинаций для каждой группы. Обе группы обладают свойством 3 – коды группы $A^{(1)}(8,2)$ и коды группы $A^{(2)}(8,2)$ отличаются в совокупности 2-мя разрядами.

Свойство 1.4. Если в группу кодов $C(n,k)$ входят прямой и инверсный коды $C_\nu = (c_\nu^1 \dots c_\nu^n) \in C(n,k)$ и $\overline{C_\nu} = (c_\nu^{-1} \dots c_\nu^{-n}) \in C(n,k)$, то в группу $A(2n,k) = C(n,k) \cdot C(n,k)$, в которую входит код $A_i = (a_i^1 \dots a_i^n) \in A(2n,k)$, может быть включен инверсный ему код $\overline{A_i} = (a_i^{-1} \dots a_i^{-n})$.

Действительно, пусть код A_i является результатом конкатенации кода $C_\nu = (c_\nu^1 c_\nu^2 \dots c_\nu^n)$ с самим собой, то есть $A_i = (c_\nu^1 c_\nu^2 \dots c_\nu^n c_\nu^1 c_\nu^2 \dots c_\nu^n)$. Но тогда, поскольку группа $C(n,k)$ содержит и прямой и инверсный коды, код $\overline{A_i} = (\overline{c_\nu^1}\overline{c_\nu^2} \dots \overline{c_\nu^n}\overline{c_\nu^1}\overline{c_\nu^2} \dots \overline{c_\nu^n})$, который образован конкатенацией кода $\overline{C_\nu}$ с самим собой $\overline{C_\nu} = (c_\nu^{-1} \dots c_\nu^{-n}) \in C(n,k)$, так же должен входить в группу $A(2n,k) = C(n,k) \cdot C(n,k)$.

Обе группы кодов, $A^{(1)}(8,2)$ и $A^{(2)}(8,2)$ (см. рисунок 5) состоят из пар прямых и инверсных кодов. Например, первый и последний, второй и предпоследний и т. д. коды обеих групп представляют собой инверсные коды.

Свойство 1.5. Если сложение по модулю 2 двух произвольных кодов:

$$C_\nu = (c_\nu^1 \dots c_\nu^n) \in C(n,k),$$
$$C_w = (c_w^1 \dots c_w^n) \in C(n,k)$$

дает код $\tilde{C} = C_\nu \oplus C_w$, который так же входит в группу $\tilde{C} \in C(n,k)$, то и в $A(2n,k) = C(n,k) \cdot C(n,k)$, в которую входят коды:

$$A_i = (a_i^1 \dots a_i^n) \in A(2n,k),$$
$$A_j = (a_j^1 \dots a_j^n) \in A(2n,k),$$

может быть включен код $\tilde{A} = A_i \oplus A_j$.

Справедливость этого утверждения следует из того, что сложение по модулю 2 левых и правых половин кодов A_i и A_j даст коды, принадлежащие группе $C(n,k)$. Следовательно, левая и правая половины кода $\tilde{A} = A_i \oplus A_j$ входят в группу кодов $C(n,k)$, а сам код $\tilde{A}$, по построению, – в группу кодов $A(2n,k)$. Это свойство дает возможность дополнительной проверки принадлежности кода к группе кодов (не только по коэффициенту корреляции). Например, для группы кодов $\Psi^{(1)}(4,2)$ (см. рисунок 3) свойство 5 выполняется, значит, будет выполняться и для всех других групп кодов, построенных на основе $\Psi^{(1)}(4,2)$ при помощи алгоритма (см. рисунок 4).

Свойство 1.6. Если код C_ν одновременно входит в две полные и максимальные группы кодов, отличающихся в совокупности на k и на $k/2$ разрядов:

$$C_\nu \in \Psi^{(g)}(n,k) \text{ и } C_\nu \in \Psi^{(d)}(n, k/2),$$

то и вся группа кодов с отличием в k разрядов принадлежит группе кодов с отличием в $k/2$ разрядов:

$$\Psi^{(g)}(n,k) \in \Psi^{(d)}\left(n, k/2\right).$$

Это свойство очевидно, т. к. для всех кодов группы $C_\nu \in \Psi^{(g)}(n,k)$ принадлежность к группе $C_\nu \in \Psi^{(d)}(n, k/2)$ является более мягким требованием.

Алгоритм 2 формирования групп кодов в полные и максимальные группы

Алгоритм объединения групп кодов, построенных при помощи алгоритма 1, в полные и максимальные группы $\Psi^{(f)}(2n,k)$ показан на рисунке 6. Группы кодов $\Psi^{(f)}(2n,k)$, которые были получены в ходе выполнения алгоритма 2, обладают определенными свойствами.

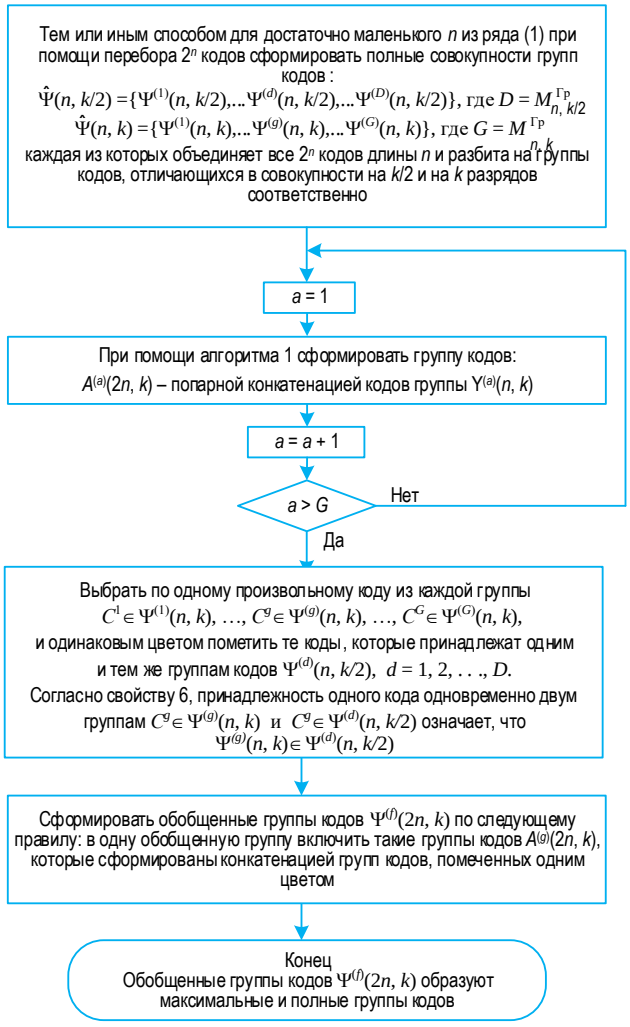


Рис. 6. Алгоритм объединения групп кодов в полные и максимальные группы

Fig. 6. An Algorithm for Combining Groups of Codes into Full and Maximum Groups

Свойство 2.1. Коды из группы $\Psi^{(f)}(2n,k)$ отличаются в совокупности не менее чем на k разрядов. Чтобы это доказать, выберем два произвольных кода $A_i \in \Psi^{(f)}(2n,k)$ и $A_j \in \Psi^{(f)}(2n,k)$.

Возможны две ситуации:

1) если коды принадлежат одной и той же группе $A_i, A_j \in A^{(g)}(2n,k)$, то они отличаются не менее, чем на k разрядов, в соответствии со свойством 3;

2) если коды принадлежат разным множествам $A_i \in A^{(g)}(2n,k)$, $A_j \in A^{(q)}(2n,k)$, то, в соответствии с алгоритмом 2, и левые, и правые их половины будут принадлежать множествам, которые отличаются на $k/2$ разрядов, а, значит, коды в целом отличаются не менее, чем на k разрядов.

Рассмотрим пример. Группы кодов, полученные выше – $\Psi^{(1)}(4,2)$ и $\Psi^{(2)}(4,2)$ (см. рисунок 3) имеют длину $n = 4$ и отличаются в совокупности на $k = 2$ разрядов. Можно заметить, что каждый код группы $\Psi^{(1)}(4,2)$ отличается от каждого кода группы $\Psi^{(2)}(4,2)$ не менее, чем на $k/2 = 1$ разрядов. Это значит, что группы кодов $A^{(1)}(8,2)$ и $A^{(2)}(8,2)$ (см. рисунок 5), у которых $2n = 8$, $k = 2$, могут быть объединены в общую группу мощностью 128 кодов: $\Psi^{(1)}(8,2) = A^{(1)}(8,2) \cup A^{(2)}(8,2)$.

Свойство 2.2. Группы кодов $\Psi^{(f)}(2n,k)$ являются полными и максимальными. Это свойство следует из полноты и максимальности групп кодов $\Psi^{(g)}(n,k)$ при $g = 1,2,\ldots G$ и полноты совокупности групп кодов:

$$\hat{\Psi}(n,k) = \{\Psi^{(1)}(n,k), \ldots \Psi^{(g)}(n,k), \ldots, \Psi^{(G)}(n,k)\}.$$

Предположим, что группа кодов $\Psi^{(f)}(2n,k)$ не полная и существует код:

$$\tilde{A} = (\tilde{a}^1 \ldots \tilde{a}^n \tilde{a}^{n+1} \ldots \tilde{a}^{2n}) \neq A_i = (a_i^1 \ldots a_i^n a_j^1 \ldots a_j^n),$$
$$A_i \in A(2n,k),$$

который может пополнить эту группу. Но тогда его левая и правая половины должны отличаться от произвольно выбранного кода $A_i \in A(2n,k)$, который сформирован конкатенацией кодов из групп $\Psi^{(1)}(n,k)$, ..., $\Psi^{(g)}(n,k)$,..., $\Psi^{(G)}(n,k)$, а этого не может быть в силу полноты последних.

Свойство 2.3. Количество кодов $M^{\text{Код}}_{n,k}$ в полной и максимальной группе кодов $\Psi^{(f)}(2n,k)$ при любом $f = 1,2,\ldots, M^{\text{Гр}}_{n,k}$ равно:

$$M^{\text{Код}}_{2n,k} = (M^{\text{Код}}_{n,k})^2 \cdot M^{\text{Гр}}_{n,k/2}. \quad (4)$$

Независимость мощности групп кодов от номера группы следует из симметричности всех процедур и алгоритмов. Формула (4) позволяет найти количество кодов в группе $\Psi^{(f)}(2n,k)$, поскольку после выполнения алгоритма 1 согласно свойству 2 в данную группу включаются $(M^{\text{Код}}_{n,k})^2$ кодов некоторой группы $A(2n,k)$, а после выполнения алгоритма 2 осуществляется объединение кодов из $M^{\text{Гр}}_{n,k/2}$ таких групп.

Для расчета числа полных и максимальных групп кодов длиной n, отличающихся в совокупности не менее чем на k, можно воспользоваться следующей формулой:

$$M^{\text{Гр}}_{n,k} = \frac{2^n}{M^{\text{Код}}_{n,k}}. \quad (5)$$

В таблице 1 показаны зависимости числа полных и максимальных групп кодов длиной n от требуемого значения k. С помощью таблицы 1 можно подобрать, с учетом требуемого коэффициента корреляции (определяется значениями n и k, см. рисунок 1), параметры группы кодов, подходящей для маркировки заданного количества объектов.

ТАБЛИЦА 1. Количество полных и максимальных групп кодов длиной *n* от требуемого числа *k*

TABLE 1. The Number of Complete and Maximum Groups of Length Codes n from the Required Number k

n – длина кода, бит	k – количество несовпадающих по значению разрядов, бит не менее			
	2	4	8	16
4	2 группы по 2^3 кодов	–	–	–
8	2 группы по 2^7 кодов	2^4 группы по 2^4 кодов	–	–
16	5 групп по 2^{11} кодов	2^8 группы по 2^{11} кодов	2^{11} группы по 2^5 кодов	–
32	–	2^{16} группы по 2^{16} кодов	2^{22} группы по 2^{10} кодов	2^{26} группы по 2^6 кодов
64	–	2^{32} группы по 2^{32} кодов	2^{44} группы по 2^{20} кодов	2^{52} группы по 2^{12} кодов
128	–	–	2^{88} группы по 2^{40} кодов	2^{104} группы по 2^{24} кодов

Алгоритм 3 формирования полной совокупности групп кодов

Алгоритм формирования полной совокупности групп кодов:

$$\hat{\Psi}(2n,k) = \{\Psi^{(1)}(2n,k), \dots \Psi^{(f)}(2n,k), \dots, \Psi^{(F)}(2n,k)\},$$

где $F = M^{\text{Гр}}_{2n,k}$, показан на рисунке 7.

Свойство 3.1. Полученная в ходе выполнения алгоритма 3 группа кодов $\Psi^{(f)}(2n,k)$ является новой полной и максимальной группой кодов, которые отличаются в совокупности на k разрядов.

Группа $\Psi^{(f)}(2n,k)$ является новой, так как выбирается код, не принадлежащий уже построенным группам кодов. Кроме того, она является полной и максимальной, поскольку операция сложения с кодом $\tilde{A}$ на последнем шаге алгоритма не может уменьшить число кодов в группе, а исходная группа $\Psi^{(1)}(2n,k)$ является полной и максимальной.

Группа кодов $\Psi^{(f)}(2n,k)$, также как и группа кодов $\Psi^{(1)}(2n,k)$, отличается в совокупности на k разрядов, т. к. сложение, выполняемое на последнем шаге алгоритма 3, приводит к инверсии разрядов одного и того же столбца матрицы $\|\Psi^{(1)}(2n,k)\|$.

Свойство 3.2. Совокупность групп кодов:

$$\hat{\Psi}(2n,k) = \{\Psi^{(1)}(2n,k), \dots \Psi^{(f)}(2n,k), \dots, \Psi^{(F)}(2n,k)\},$$

где $F = M^{\text{Гр}}_{2n,k}$, полученная в результате выполнения алгоритма 3, является полной. Это свойство следует из того, что на 3 шаге алгоритма 3 выбираются все коды, которые еще не были включены в существующие группы кодов.

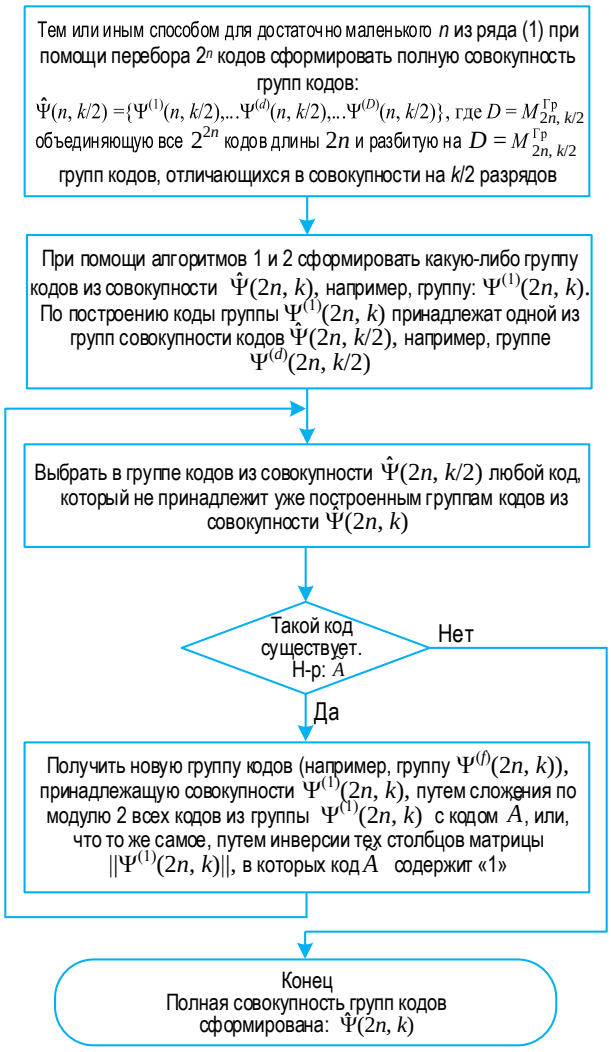


Рис. 7. Алгоритм формирования полной совокупности групп кодов

Fig. 7. The Algorithm for the Complete Set Formation of Code Groups

Заключение

Радиочастотная идентификация объектов относится к числу динамично развивающихся технологий. Большая защищенность от подделок, способность работать при любой освещенности, различных погодных и климатических условиях, надежность регистрации маркированных объектов определили потребность в подобных системах в таких областях, как почтовая связь, логистика, транспорт, медицина, коммерческая деятельность, оптовая и розничная торговля, охранные системы, таможенная деятельность и многих других [1–3]. С информационной точки зрения работа таких систем невозможна без использования процедур синтеза групп кодов для

маркировки того или иного множества объектов. Причем к свойствам кодов и групп кодов предъявляются дополнительные жесткие требования, обусловленные необходимостью маркировки и надежной регистрации большого числа объектов в условиях неидеальной среды передачи радиосигналов [14].

При создании RFID-систем возникает проблема множественного доступа к группе идентифицируемых объектов, расположенных в ограниченном пространстве. Решение этой проблемы предложено в стандартах GEN1 и GEN2 EPC Global применительно к активным RFID-меткам, построенным на кремниевой технологии, использующим автономные источники питания и имеющим относительно высокую стоимость. В этом случае множественный доступ реализуется путем информационного воздействия на метку при помощи специальных команд типа: «Сообщи заводской номер», «Передай (запиши) байт из (в) память с адресом А», «Задержи передачу на время *t*», «Перейди (выйди) в (из) режима молчания» и т. п. Примерами таких алгоритмов могут служить бинарный алгоритм, *Q*-алгоритм (или алгоритм запроса с параметром *Q* С. Смита), алгоритм побитового перебора и др. [15].

В работе сформулирована задача множественного доступа для системы идентификации объектов при помощи применения пассивных RFID-меток на ПАВ, не использующих источники питания и имеющих низкую стоимость.

В ходе исследования был предложен подход к формированию групп слабо коррелированных кодов для маркировки объектов, расположенных в ограниченном пространстве. Разработан комплекс алгоритмов, позволяющий на основе корреляционного анализа формировать для меток группы кодов с заданными длиной кода, мощностью и коэффициентом попарной корреляции. Приведено обоснование свойств синтезируемых групп кодов.

Список источников

1. Колбанёв М.О., Верзун Н.А. Цифровая трансформация логистических процессов множественной идентификации объектов // Развитие науки и научно-образовательного трансфера логистики / под научной ред. д-ра экон. наук, проф. В.В. Щербакова. СПб.: Санкт-Петербургский государственный экономический университет, 2019. С. 53–69. EDN:ULRAZJ

2. Корочкин Л.С., Астафьев И.А., Молдованов А.А., Шмаков М.С. Радиочастотная идентификация товарно-транспортных накладных грузов // Труды БГТУ. Серия 4: Принт- и медиатехнологии. 2019. № 1(219). С. 24–28. EDN:VIDAOG

3. Верзун Н.А., Колбанёв М.О., Омельян А.В. RFID-технологии для эффективности и безопасности документооборота // Технологии информационно-экономической безопасности. СПб.: Санкт-Петербургский государственный экономический университет, 2016. С. 44–51. EDN:XFZLDH

4. ГОСТ Р ИСО/МЭК 15693-3-2011. Карты идентификационные. Карты на интегральных схемах бесконтактные. Карты удаленного действия. Антиколлизия и протокол передачи данных. М.: Стандартинформ, 2014. 40 с.

5. Liu H., Chen Y., Tzeng W. A Multi-Carrier UHF Passive RFID System // Proceedings of the International Symposium on Applications and the Internet Workshops (Hiroshima, Japan, 15–19 January 2007). IEEE, 2007. DOI:10.1109/SAINT-W.2007.9

6. Верзун Н.А., Воробьева Д.М., Колбанёв А.М., Колбанёв М.О. Обзор технологий и стандартов RFID систем // Информационные технологии и телекоммуникации. 2018. Т. 6. № 1. С. 1–11. EDN:XOCWYX

7. Лахири С. RFID. Руководство по внедрению: практическое руководство от опытного ИТ-архитектора в области RFID: научитесь оценивать, планировать и развертывать RFID-системы. Пер. с англ. М.: КУДИЦ-ПРЕСС, 2007. 298 с.

8. Profetto L., Gherardelli M., Iadanza E. Radio Frequency Identification (RFID) in health care: where are we? A scoping review // Health and Technology. 2022. Vol. 12. PP. 879–891. DOI:10.1007/s12553-022-00696-1

9. Hubacz M., Pawłowicz B., Salach M., Trybus B. Model urządzenia piorącego wykorzystującego tekstroniczne transpondery RFID // Pomiary Automatyka Robotyka. 2022. Vol. 26. No. 4. PP. 69–77.

10. Вековцева Т.А., Шанина Т.В. Технология RFID и будущее производство радиочастотной этикетки // Международный научно-исследовательский журнал. 2017. № 3-4(57). С. 20–22. DOI:10.23670/IRJ.2017.57.071. EDN:YGTZHX

11. Морелос-Сарагоса Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение. М.: Техносфера, 2005. 319 с.

12. Березкин Е.Ф. Основы теории информации и кодирования: учебное пособие. СПб.: Лань, 2022. 320 с.

13. Антонов В.В. Антиколлизионный механизм информационного обмена для маркеров на поверхностно-активных волнах // VI Международная научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (Санкт-Петербург, Российская Федерация, 01–02 марта 2017 г.). СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2017. Т. 3. С. 36–39. EDN:YTBDSU

14. Шарфельд Т. Системы RFID низкой стоимости. М., 2006. 197 с.

15. Верзун Н.А., Колбанёв А.М., Советов Б.Я., Колбанёв М.О. Антиколлизионные алгоритмы систем радиочастотной идентификации // Известия СПбГЭТУ ЛЭТИ. 2018. № 10. С 24–31. EDN:PLTCCZ

References

1. Kolbanev M.O., Verzun N.A. Digital transformation of logistics processes for multiple identification of objects. *Development of Science and Scientific-Educational Transfer of Logistics* / under the scientific editorship of Dr. of Economics, prof. V.V. Shcherbakov. St. Petersburg: St. Petersburg State University of Economics Publ.; 2019. p.53–69 (in Russ.) EDN:ULRAZJ

2. Korochkin L.S., Astafyev I.A., Moldovanov A.A., Shmakov M.S. Radio frequency identification of consignment notes. *Proceedings of BSTU. Series 4: Print and media technologies*. 2019;1(219):24–28. (in Russ.) EDN:VIDAOG
3. Verzun N.A., Kolbanev M.O., Omelyan A.V. RFID technologies for the efficiency and security of document flow. *Technologies of Information and Economic Security*. St. Petersburg: St. Petersburg State University of Economics Publ.; 2016. p.44–51. (in Russ.) EDN:XFZLDH
4. GOST ISO/IEC 15693-3-2011. *Identification cards. Contactless integrated circuit cards. Vicinity cards. Part 3. Anticollision and transmission protocol (IDT)*. Moscow: Standartinform Publ.; 2014. 40 p. (in Russ.)
5. Liu H., Chen Y., Tzeng W. A Multi-Carrier UHF Passive RFID System. *Proceedings of the International Symposium on Applications and the Internet Workshops, 15–19 January 2007, Hiroshima, Japan*. IEEE; 2007. DOI:10.1109/SAINT-W.2007.9
6. Verzun N., Vorobyova D., Kolbanev A., Kolbanev M. Review of Technologies and Standards RFID System. *Telecom IT.* 2018;6(1):1–11. EDN:XOCWYX
7. Lahiri S. *RFID Sourcebook Upper Saddle River*. IBM Press, 2006. 276 p.
8. Profetto L., Gherardelli M., Iadanza E. Radio Frequency Identification (RFID) in health care: where are we? A scoping review. *Health and Technology*. 2022;12:879–891. DOI:10.1007/s12553-022-00696-1
9. Hubacz M., Pawłowicz B., Salach M., Trybus B. Model of Washing Device Using Textronic RFID Transponders. *Pomiary Automatyka Robotyka*. 2022;26(4):69–77.
10. Vekovtseva T.A. Shanina T.V. RFID technology and future production of radio frequency labels. *International Research Journal.* 2017;3-4(57):20–22. DOI:10.23670/IRJ.2017.57.071. EDN:YGTZHX
11. Morelos-Zaragoza R. *The Art of Noise-Tolerant Coding. Methods, Algorithms, Application*. Moscow: Tehnosfera Publ.; 2005. 319 p. (in Russ.)
12. Berezkin E.F. *Fundamentals of Information Theory and Coding*. St. Petersburg: Lan Publ.; 2022. 320 p. (in Russ.)
13. Antonov V. Anti-Collision Mechanism of Information Exchange for Markers on Surface-Active Waves. *Proceedings of the IVth International* Conference *on Infotelecommunications in Science and Education, 01–02 March 2017, St. Petersburg, Russian Federation, vol.3.* St. Petersburg: The Bonch-Bruevich Saint-Petersburg State University of Telecommunications Publ.; 2017. p.36–39. (in Russ.) EDN:YTBDSU
14. Sharfeld T. *Low Cost RFID Systems*. Moscow; 2006. 197 p. (in Russ.)
15. Verzun N.A., Kolbanev A.M., Sovetov B.Ya., Kolbanev M.O. Anti-collision algorithms of radio-frequency identification systems. *Proceedings of Saint Petersburg Electrotechnical University.* 2018;10:24–31. (in Russ.) EDN:PLTCCZ

Статья поступила в редакцию 15.08.2024; одобрена после рецензирования 21.11.2024; принята к публикации 28.11.2024.
The article was submitted 15.08.2024; approved after reviewing 21.11.2024; accepted for publication 28.11.2024.

Информация об авторах:

ВЕРЗУН Наталья Аркадьевна — кандидат технических наук, доцент, доцент кафедры информационных систем и технологий Санкт-Петербургского государственного экономического университета
https://orcid.org/0000-0002-0126-2358

КОЛБАНЁВ Алексей Михайлович — технический руководитель продукта Новостройки направления по работе с девелоперами АО «ЭР-Телеком Холдинг»
https://orcid.org/0009-0008-7542-9123

КОЛБАНЁВ Михаил Олегович — доктор технических наук, профессор, профессор кафедры информационных систем и технологий Санкт-Петербургского государственного экономического университета
https://orcid.org/0000-0003-4825-6972

Авторы сообщают об отсутствии конфликтов интересов.
The authors declare no conflicts of interests.

Обзорная статья
УДК 004.056(075.58)
https://doi.org/10.31854/1813-324X-2024-10-6-79-98

Прогресс в теории прикладной криптографии: обзор и некоторые новые результаты. Часть 2. Бесключевая криптография

Валерий Иванович Коржик[1] ✉, val-korzhik@yandex.ru
Виктор Алексеевич Яковлев[1], yakovlev.va@sut.ru
Владимир Сергеевич Старостин[1], vm.ffp@sut.ru
Михаил Викторович Буйневич[2], bmv1958@yandex.ru

[1]Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация
[2]Санкт-Петербургский университет ГПС МЧС России, Санкт-Петербург, 196105, Российская Федерация

Аннотация

Настоящая работа является второй частью статьи «Прогресс в теории прикладной криптографии: обзор и некоторые новые результаты», опубликованной в четвертом номере журнала ТУЗС за 2024 год. Она посвящена специфическому разделу так называемой бесключевой криптографии (БК). Актуальность данной статьи состоит в том, что рассматриваемые в ней методы позволяют обеспечить конфиденциальность передачи информации по открытым каналам связи, либо вообще не выполняя никакого предварительного ее шифрования, а эксплуатируя лишь преобразования, естественно происходящие в каналах связи, либо, применяя обычную (ключевую) криптографию, ключи для которой передаются по открытым каналам связи с использованием методов БК.

Настоящая работа начинается с описания вайнеровской концепции подслушивающего канала и методов кодирования в нем, обеспечивающих надежную передачу по основному каналу с гарантированным малым количеством шенноновской информации, утекающей по каналу перехвата. Далее исследуются сценарий с коммутативной криптографией (CE) и протокол, обеспечивающий конфиденциальность передачи информации без всякого обмена ключами. Следующая модель относится к многолучевому каналу и применению MIMO-технологии по протоколу Дина и Голдсмит. Доказывается, что для него секретность передачи обеспечивается только при ограничении на количество приемных антенн перехватчика.

Следующий сценарий использует технологию антенн с управляемой диаграммой (VDA), причем устанавливаются условия на многолучевость и расположение корреспондентов радиосвязи, при которых может быть обеспечена конфиденциальность передачи информации.

Анализируется также недавно предложенная криптосистема EVESkey. Доказывается, что существует простая атака, которая может нарушить ее конфиденциальность.

Описывается ряд протоколов, выполняемых по бесшумному открытому каналу, которые, однако, не являются стойкими, поскольку они имеют нулевую секретную пропускную способность. Доказывается, что при матричном обмене в канале связи типа Интернет, конфиденциальность может быть обеспечена только по критерию ограничения сложности декодирования. В конце работы формулируются фундаментальные проблемы прикладной криптографии, решение которых могло бы значительно стимулировать дальнейшее развитие этой отрасли науки.

Ключевые слова: *бесключевая криптография, отводной канал, MIMO-технология, физический уровень секретности, усиление секретности, коммутативное шифрование.*

Ссылка для цитирования: Коржик В.И., Яковлев В.А., Старостин В.С., Буйневич М.В. Прогресс в теории прикладной криптографии: обзор и некоторые новые результаты. Часть 2. Бесключевая криптография // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 79–98. DOI:10.31854/1813-324X-2024-10-6-79-98. EDN:HPBOWG

Review research
https://doi.org/10.31854/1813-324X-2024-10-6-79-98

Advance in Applied Cryptography Theory: Survey and Some New Results. Part 2. Keyless Cryptography

Valery I. Korzhik[1] ✉, val-korzhik@yandex.ru
Viktor A. Yakovlev [1], yakovlev.va@sut.ru
Vladimir S. Starostin[1], vm.ffp@sut.ru
Mikhail V. Buinevich[2], bmv1958@yandex.ru

[1]The Bonch-Bruevich Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation
[2]Saint Petersburg University of State Fire Service of Emercom of Russia,
St. Petersburg, 196105, Russian Federation

Annotation

Actuality. *The current paper is the second part of the paper "Advance in Applied Cryptography Theory: Survey and Some New Results. Part 1. Key Cryptography" published in the journal PTU, n.4, 2024. It is devoted to such specific area of applied cryptography as keyless one (KC\ Actuality of the current paper consists in the fact that considered in it methods allow to provide a confidentiality of information transmission over public communication channels, either without any its encryption in advance, executing a natural properties of communication channels or executing conventional key cryptography but with the keys which are elaborated before by means of KC.*
The natural properties of communication channels can be the following: additive noise, multiray wave propagation, MIMO technology and existence of feedback channel.
Our paper starts with a consideration of Wyner's concept of wire-tap channels and corresponding to it encoding and decoding methods providing very reliable information transmission over the main channels and negligible amount of information leaking over the wire-tap channels to eavesdroppers. Next it is investigated scenario with a commutative encryption (CE) and corresponding protocol of message exchange over ordinary noiseless public channel that provides security of encrypted information but without any key exchange between users in advance. It is proved which of well known symmetric and asymmetric ciphers are commutative or non-commutative ones. Next model concerns a fading channels under the application of Dean-Goldsmith protocol in frames of MIMO technology. We are proving that this protocol is secure if, and only if, the number of eavesdropper antennas is less than the number of antennas at legitimate users. Next scenario executes variable directional antennas (VDA) and it is proved for which conditions on a locations of legitimate users and eavesdroppers such approach occurs secure given the number of propagation rays is at least two.. We show in the next chapter that there is an attack compromising of recently proposed EVESkey cryptosystem and hence such one is not secure in spite of the statement of its authors.
Finally, we investigate several protocols intended for key sharing over noiseless constant public channels (like Internet) and established that they are mostly insecure because have all zero secret capacity. Only one protocol based on matrix channel exchange is able to provide security of key sharing but in terms of the required breaking complexity. Thus such approach can be used only for the case when legitimate users belong to low level of security requirements.
At the end of the paper we formulate several fundamental problems of applied cryptography which after of their solutions could be very useful for practice.

Keywords: *keyless cryptography, wire-tap channel concept, physical level security, MIMO technology, privacy amplification, commutative encryption*

For citation: Korzhik V.I., Yakovlev V.A., Starostin V.S., Buinevich M.V. Advance in Applied Cryptography Theory: Survey and Some New Results. Part 2. Keyless Cryptography. *Proceedings of Telecommunication Universities.* 2024;10(6):79–98. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-79-98. EDN:HPBOWG

1. Введение

В части 1 настоящей статьи уже отмечалось, что термин «бесключевая криптография» (*от англ.* Keyless Cryptography) был предложен еще в 1983 г. в статье Б. Альперна и Ф. Шнайдера [1]. В последующие годы появилось множество работ, посвящен-

ных этой тематике, например [2–5] и другие. Попытаемся дать достаточно общее определение этому виду прикладной криптографии, хотя нужно признать, что термин «бесключевая криптография» не стал повседневно используемым, и иногда его заменяют частными случаями из данной области, например, «обеспечение секретности на физическом уровне» [5], сценарием с «отводным каналом» [6] и даже «квантовой криптографией» [7].

Само же общее понятие «бесключевой криптографии» можно определить следующим образом: это совокупность методов обеспечения секретности (конфиденциальности) передачи или хранения представленной в цифровом виде информации, в том числе и цифровых ключевых данных, без использования методов шифрования / дешифрования любого вида из арсенала методов современной криптографии или без предварительного обмена ключами для традиционной криптографии.

На первый взгляд, такое определение похоже на «оксюморон», поскольку криптография, казалось бы, основана именно на использовании различных преобразований открытой информации и секретных ключей, доступных только легитимным (то есть уполномоченным) пользователям. Даже несимметричная (или двухклассная криптография, или иначе – криптография с открытыми ключами (PKC, *аббр. от англ*. Public Key Cryptography), предложенная У. Диффи и М. Хеллманом, хотя и явившаяся подлинной революцией в криптографии, требует использования пары ключей, один из которых (шифрования), не должен быть секретным, но его необходимо передавать по каналам связи всем пользователям. Кажущееся противоречие в определении «бесключевая криптография» объясняется достаточно просто: секретность обеспечивается, прежде всего, преобразованиями, производимыми в каналах связи, между источниками информации и их получателями. В этом смысле так называемая квантовая криптография (QC, *аббр. от англ.* Quantum Cryptography) представляла бы собой частный случай бесключевой, поскольку ее секретность обеспечивается законами квантовой физики (принцип Гейзенберга, парадокс Эйнштейна – Подольского – Розена и т. д.). Однако, в настоящей статье мы не будем описывать направление QC и не столько потому, что авторы не являются специалистами в этой области. Так, один из них еще в конце прошлого века делал заказной доклад в лаборатории Calderonlab Оксфордского университета (Англия) на тему «Протоколы квантовой криптографии»; он же был научным руководителем кандидатской диссертации [8], посвященной именно исследованиям в области QC. Главная же причина того, что в данную статью не включена QC, состоит в том, что это специальная область знаний, где значительно большее внимание уделяется даже не прикладной математике, а физике (см. например, упомянутую выше диссертацию [8] и учебник «Основы криптографии», 3-е издание [9]). Кроме того, сейчас опубликовано много работ по QC (см., например [7, 10, 11] и др.), к которым мы не сможем добавить какие-либо свои новые результаты. Наконец, линии связи для передачи ключевых данных, протяженностью до 1000 км сейчас реализованы практически и достаточно хорошо описаны в [11]. Заметим также, что, если цель бесключевых криптосистем состоит в защите от перехвата самих ключевых данных, передаваемых по каналам связи, то на втором этапе ее выполнения могут использоваться общие (как симметричные, так и несимметричные) криптосистемы.

В работе [12] приводится обзор криптосистем, обеспечивающих секретность на физическом уровне, и также используется термин бесключевая криптография. Причем отмечается, что данные методы обеспечения конфиденциальности являются перспективными для технологии 6G.

Настоящая статья структурирована следующим образом. В разделе 2 описана бесключевая криптография, основанная на модели, когда в каналах связи, по которым предполагается передавать конфиденциальную информацию, присутствуют внешние шумы. Такой сценарий носит название Вайнеровской модели канала с отводом. Для данного сценария бесключевая криптография оказывается особенно эффективной в случае необходимости защиты от утечки информации по так называемым побочным каналам. В разделе 3 предлагается изучить сценарий передачи информации, когда между корреспондентами имеются бесшумные дуплексные каналы связи и используются так называемые коммутативные алгоритмы шифрования. В разделе 4 описываются сценарии, где имеются многолучевые (замирающие) каналы связи между корреспондентами, и по существу секретность обмена информацией обеспечивается различием некоторых характеристик таких каналов. В разделе 5 приведены новые результаты по предложенному японскими авторами методу распределения ключей при наличии смарт-антенны с управляемой диаграммой направленности и многолучевого канала связи между корреспондентами. Такой подход в дальнейшем был развит в работе авторов настоящей статьи. Раздел 6 иллюстрирует ситуацию, когда кажущаяся секретной бесключевая криптосистема на самом деле таковой не является. Это делается на примере EVESkey-схемы, сравнительно недавно предложенной китайскими учеными. Раздел 7 относится к методам передачи конфиденциальных ключевых данных по бесшумным дуплексным каналам с обратной связью типа Интернет. Раздел 8 подытоживает основные результаты настоящей работы.

2. Сценарий Вайнеровской концепции шумных каналов с отводом

Статья А. Вайнера [6] определила важное направление в теории, а, возможно, и на практике информационной безопасности. По нашему мнению, автора этой статьи А. Вайнера можно было бы смело назвать «Шенноном информационной безопасности». Действительно, К.Э. Шеннон в своей фундаментальной работе [13] впервые строго доказал, что существуют сколь угодно надежные методы передачи информации по каналам с помехами, причем достижимые не за счет тривиального уменьшения скорости передачи или увеличения отношения мощности сигнала к мощности шума, а за счет специальной обработки сигналов на передаче и приеме, т. е. кодирования и декодирования; к тому же существует максимальная скорость передачи, названная Шенноном *пропускной способностью канала связи.*

А. Вайнер доказал, что в каналах с отводом (т. е. с возможностью перехвата информации «оппонентами») существует такой метод кодирования / декодирования, который обеспечивает сколь угодную достоверность передачи информации по основному (легальному каналу) и сколь угодно малую утечку информации по отводному каналу. Им была также рассчитана максимально возможная при этом скорость передачи по основному каналу, названная им *секретной пропускной способностью.*

На рисунке 1 представлена блок-схема шумного канала связи с отводом и кодированием / декодированием по Вайнеру.

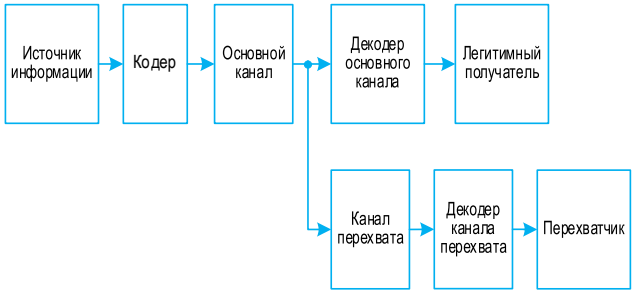


Рис. 1. Блок схема канала с отводом и кодированием / декодированием по Вайнеру

Fig. 1. Block-Scheme of Wire-Tap Channel with of Encoder / Decoder Due to Wyner

Основной особенностью канала, показанного на рисунке 1, является то, что канал перехвата всегда оказывается более шумным, чем основной канал. Поэтому его называют иногда *вырожденным каналом перехвата*.

На рисунке 2 представлена блок-схема Вайнеровской концепции с каналом перехвата, статистически независимым от основного канала, но с шумом, которая более приемлема для практики.

Не умаляя общности любых симметричных каналов без памяти в качестве основного и отводного каналов, рассмотрим далее двоичные симметрические каналы (ДСК), как из соображений большей простоты их рассмотрения, так и из-за частости их практического присутствия.

В работе [14], вскоре последовавшей после «пионерской» работы А. Вайнера [6], было доказано, что секретная пропускная способность C_s канала с отводом, показанном на рисунке 2, задается следующей формулой:

$$C_s = C_m - C_w, \qquad (1)$$

где C_m – шенноновская пропускная способность основного ДСК с вероятностью ошибки символа P_m: $C_m = 1 + P_m \log_2 P_m + (1 - P_m)\log_2 (1 - P_m)$; C_w – шенноновская пропускная способность отводного ДСК с вероятностью ошибок символа $P_w > P_m$: $C_w = 1 + P_w \log_2 P_w + (1 - P_w) \log_2 (1 - P_w)$.

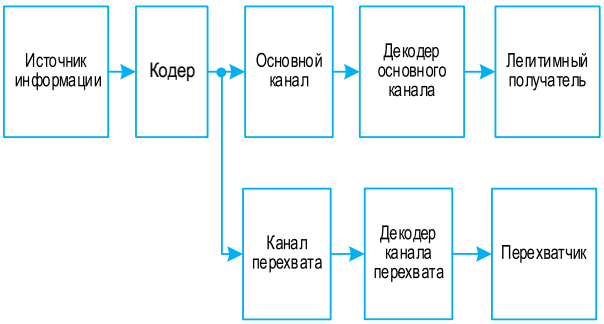


Рис. 2. Блок-схема с каналом перехвата, независимого по шумам от основного канала

Fig. 2. Block Scheme of Interception Channel Noisy Independent on the Main Channel

Если же отводной канал оказывается менее шумным, то есть $P_w < P_m$, то C_s нужно рассчитывать C_s по следующей формуле [14]:

$$C_s = h(P_m + P_w - 2P_m * P_w) - h(P_w), \qquad (2)$$

где $h(x) = x \log_2 x + (1 - x) \log_2 (1 - x)$.

При рассмотрении побочных каналов утечки информации (гальванических, оптических, вибрационных, электромагнитных) типичным является условие $P_w >> P_m$ и даже $P_m = 0$ (например, утечка по электромагнитным каналам от соединительных кабелей между выходом дешифратора и принтерами или мониторами). В этом случае формула (1) преобразуется к виду:

$$C_s = 1 - C_w. \qquad (3)$$

Важным отличием кодирования при помощи кодов с исправлением ошибок [16] от кодирования для концепции канала с отводом является то обстоятельство, что в последнем случае требуется рандомизация при помощи дополнительного шума. Такое кодирование в работе [15] было названо *кодовым зашумлением* (КЗ), которое может быть до-

статочно просто реализовано при помощи использования линейных блочных кодов и шумового генератора.

Пусть S – это двоичная информационная цепочка длиной k, подлежащая защите от перехвата при помощи технологии КЗ; γ – чисто случайная двоичная цепочка символов длиной $n-k$, получаемая от физического генератора шума; $V-(n,n-k)$ двоичный линейный код с длиной блока n и с числом информационных символов $n-k$; $c=f(\gamma)$ – проверочные символы кода V, полученные после кодирования цепочки γ. Тогда кодовый блок y для КЗ формируется следующим образом:

$$y=(x_1,x_2)=(\gamma, S \oplus c), \qquad (4)$$

где (x_1,x_2) – конкатенация цепочек x_1 и x_2; $\oplus$ – операция побитового сложения по mod2.

Легко видеть, что при отсутствии шума в основном канале $(P_m=0)$ сообщение $\check{S}$ восстанавливается без ошибок, т. е. $\check{S}=S$, как:

$$\check{S}= x_2 \oplus f(x_1). \qquad (5)$$

Основная теорема А. Вайнера утверждает, что в асимптотике, т. е. когда длина кодового блока $n\to\infty$ и выбран оптимальный код, максимальное количество информации, утекающее к перехватчику (т. е. пропускная способность канала утечки) $C_0=\max\limits_{s}(I(s;x)/k)$. На практике может интересовать и неасимптотика, когда $n<\infty$, и выбран какой-то конкретный линейный код V.

Точная формула для расчета этой величины C_0 была получена в работе [15]:

$$C_0=1+\frac{1}{k}\sum_{j=1}^{2^k}P(Vj)\log_2 P(Vj), \qquad (6)$$

где $P(V_j)=\sum_{i=1}^{n}N_{ij}P_w^i(1-P_w)^{n-i}$, N_{ij} – количество слов Хэммингова веса i в j-ом смежном классе стандартной расстановки V_n/V.

Расчет C_0 по формуле (6) осложняется тем обстоятельством, что коэффициенты N_{ij} (называемые обычно *спектром смежных классов кода V*) известны не для всех классов линейных кодов. Они известны для кодов: Хэмминга, БХЧ, исправляющих двухкратные ошибки, Голея, некоторых самодуальных, симплексных и кода Рида – Малера (см. подробности в работе [16]). К сожалению, расчет спектров для произвольных линейных кодов связан с экспоненциальными трудностями). Пример расчета C_0 по (6) для кода Голея – $(24,12)$, приведен в таблице 1 (см. также [17]). В последней строке этой таблицы показаны величины пропускной способности канала утечки $Co(P_w)$ без применения КЗ. Видно, что начиная с $P_w=5*10^{-2}$, использование КЗ демонстрирует его значительное преимущество.

ТАБЛИЦА 1. Результат расчета пропускной способности канала утечки C_0 при использовании в качестве КЗ линейного кода Голея – (24, 12) для различных вероятностей P_w

TABLE 1. Calculation of the Secrete Capacity C_0 for Goley Code – (24, 12) Given Different Probabliyties P_w

P_w	10^{-3}	10^{-2}	$5*10^{-2}$	10^{-1}	$2*10^{-1}$
C_0	0,977	0,838	0,458	0,156	$6{,}3*10^{-3}$
$C(P_w)$	≈ 1	0,91	0,72	0,53	0,28

Заметим, что использование КЗ для защиты от утечки по побочным каналам можно сочетать с такими традиционными средствами, как зашумление линий связи или окружающего аппаратуру пространства специальными шумовыми генераторами, экранировка кабелей и устройств, введение контролируемых зон и т. п. В этих случаях эффективность КЗ можно оценивать энергетическим выигрышем от его применения, который рассчитывается, как и для обычных систем связи:

$$\eta = 20\lg\frac{h_{\text{кз}}}{h}, \qquad (7)$$

где $h_{\text{кз}}$, h – отношение сигнал / шум с применением КЗ и без него, соответственно, для которых обеспечивается равная защищенность информации. В работе [17] показано, что для кодов с известными спектрами смежных классов, с защищенностью (в терминах пропускной способности) $C_0=10^{-3}$, энергетический выигрыш колеблется в пределах 20÷30 дБ, что даже превосходит энергетический выигрыш от применения некоторых кодов для обычного исправления ошибок в каналах связи.

Конечно, КЗ может использоваться не только для защиты от утечки по побочным каналам, но и – от перехвата информации по обычным каналам связи. Однако в этом случае зашумление основных каналов может оказаться даже больше, чем каналов перехвата. Для такого сценария У. Маурером был разработан специальный метод открытого обсуждения (PD, *аббр. от англ.* Public Discussion) [14]. Цель этой процедуры состоит в том, чтобы свести первоначальное условие $(P_m>P_w)$ к новому условию – $(P_m<P_w)$, но, конечно, без потери секретности передаваемой информации. Им же были предложены различные способы выполнения PD. Например, A передает к B каждый двоичный символ «S» раз, а B удерживает только такие S-блоки, в которых количество единиц (или нулей) оказывается больше некоторого заранее выбранного порога. Второй подход, также предложенный в [14], заключается в выполнении следующего интерактивного протокола между A и B.

Пусть E и D обозначают образцы ошибок в основном канале $(A\to B)$ и в канале перехвата $(A\to E)$, соответственно. Если A посылает B цепочку случайных бит X, то B принимает цепочку $Y=X+E$, а перехватчик – цепочку $Z=X+D$. Далее B посылает A

цепочку $W = Y + V$ по открытому каналу, где V – некоторая цепочка случайных бит, сгенерированных B. Наконец A вычисляет $W + X = V + E$, т. е. A принимает V с прежней битовой ошибкой P_m, тогда как E, приняв $Z = X + D$ и $W = X + E + V$, может лишь вычислить $Z + W = V + E + D$ с битовой ошибкой $P_m + P_w - 2\,P_m * P_w > P_w$.

Наконец, после выполнения процедуры PD, что обеспечивает выполнение условия $P_m < P_w$, легальные пользователи A и B могут исправить все ошибки в цепочке данных, которыми они, в конце концов, обменялись, а затем выполнить еще преобразование этих финальных данных, которое называется *усилением секретности* (PA, *аббр. от англ.* Privacy Amplification). PA представляет собой не что иное, как хеширование разделенной между A и B информации с использованием хеш-функции $h(..)$ из универсального класса, т. е.:

$$h(x) = [xxh]_k,$$

где x – умножение в конечном поле $GF(2^n)$, $x, h \in \in GF(2)^n)$; $[y]_k$ – удержание k наименьших значимых бит цепочки y.

В работе [18] доказана *обобщенная теорема усиления секретности*, которая дает верхнюю границу количества шенноновской информации, утекающей к перехватчику, после выполнения процедуры PA:

$$I_0 \leq \frac{2^{-(k-t_c-t-r)}}{\gamma \ln 2}, \qquad (8)$$

где k – полная длина цепочки, разделенная A и B после выполнения протокола PD; t_c – информация Реньи (или иначе – collision information), полученная E после перехвата цепочки x; t – финальная длина цепочки после выполнения протокола PA; r – количество проверочных символов, переданных от A к B для согласования их данных и переданных по открытому каналу; γ – коэффициент, приближающийся к 0,42, когда $k, k - t \to \infty$.

Для t_c справедлива следующая формула:

$$t_c = k(1 + \log_2(P_w{}^2 + (1 - P_w)^2)). \qquad (9)$$

Таким образом, если известна зашумленность основного (P_m) и отводного (P_w) каналов, то можно так выбрать параметры системы кодового зашумления, чтобы обеспечить утечку к перехватчику не более заданного наперед количества шенноновской информации и максимизировать, насколько это возможно, скорость передачи по основному каналу. Следует особо отметить, что перед выполнением протокола КЗ необходимо договориться между легитимными пользователями о надежном обеспечением их аутентификации. Более подробное описание протокола передачи информации в условиях активного перехвата (т. е. возможного навязывания перехватчиком ложной информации) можно найти в статье [19].

3. Коммутативное шифрование

Данный метод обеспечения информационной безопасности, при возможности дуплексного обмена информацией по каналам связи между пользователями, был впервые предложен А. Шамиром и описан в монографии [20]. Дальнейшее его исследование производилось в малодоступной для российских специалистов книге [21]. Поэтому приведем его в настоящем разделе с некоторыми «расширениями».

Пусть имеется шифр, обладающий свойством так называемого *коммутативного шифрования* (CE, *аббр. от англ.* Commutative Encryption). Заметим, что сам термин «коммутативное шифрование» не слишком укоренился в публикациях по тематике «криптография». Действительно, в почти 700-страничной книге [22] в разделе «Index» такой термин вообще отсутствует. Более того, в монографии [23] похожий термин «*некоммутативная криптография*» (non-commutative cryptography) означает совершенно другое, а именно использование некоммутативных (неабелевых) групп для построения несимметричных криптосистем.

Будем называть коммутативным шифрованием (КШ) любой алгоритм шифрования, если для него при любых ключах и любых сообщениях выполняется следующее условие:

$$f_{k_1}(f_{k2}(M)) = f_{k_2}(f_{k1}(M)), \qquad (10)$$

где $f_{k_1}(M)$ или $f_{k_2}(M)$ – алгоритм шифрования с использованием любых сообщений и любых ключей k_1, k_2.

На рисунке 3 показан протокол, который позволит пользователю A передать по бесшумному и открытому каналу связи другому пользователю B любое сообщение M, зашифрованное при помощи алгоритма КШ с выполнением следующих свойств:

– между корреспондентами A и B не требуется выполнения никакого предварительного протокола по обмену секретными ключами;

– стойкость данного протокола полностью эквивалентна стойкости исходной криптосистемы $f_k(M)$.

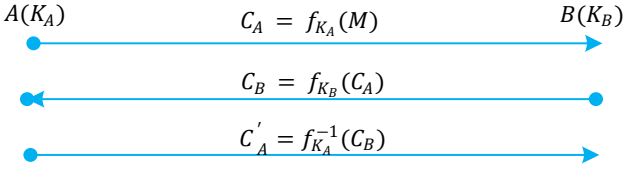


Рис. 3. Протокол передачи секретных сообщений при использовании КШ

Fig. 3. Protocol of Secret Message Transmission with the Use of CE

Действительно, первоначально A и B генерируют независимо свои ключи шифрования и дешифрования (они совпадают для симметричных криптосистем) и хранят их недоступными для посторонних пользователей. Далее они выполняют 3-шаговый протокол обмена по каналу, соединяющему A с B и B с A (т. е. такой канал должен между ними существовать). На первом шаге A передает B криптограмму, полученную им при использовании личного ключа K_A. Получив такую криптограмму C_A от A, пользователь B шифрует ее (как сообщение) своим личным ключом K_B и передает обратно пользователю A как криптограмму $C_B = f_{K_B}(C_A) = = f_{K_B}\left(f_{K_A}(M)\right)$.

Используя свойство коммутативности шифрования (10), получаем:

$$C_B = f_{K_B}\left(f_{K_A}(M)\right) = f_{K_A}\left(f_{K_B}(M)\right).$$

Тогда пользователь A может применять процедуру дешифрования на своем ключе дешифрования, соответствующему ключу шифрования, т. е. $f_{K_A}^{-1}(C_B)$.

После выполнения процедуры такого дешифрования A получает криптограмму C_A', которую и посылает B на третьем шаге протокола:

$$C_A' = f_{K_A}^{-1}(C_B) = f_{K_A}^{-1}(f_{K_A}\left(f_{K_B}(M)\right) = f_{K_B}(M).$$

Тогда B тривиально применяет к полученной криптограмме C_A' процедуру дешифрования $f_{K_B}^{-1}(C_A')$ на ключе дешифрования, соответствующему своему ключу шифрования K_B, что дает окончательное выражение:

$$f_{K_B}^{-1}(C_A') = f_{K_B}^{-1}(f_{K_B}(M)) = M,$$

что и требовалось доказать.

Однако пока остается открытым вопрос – какие из известных симметричных (или несимметричных) шифров обладают свойством коммутативности (10)? Прежде всего заметим, что такие популярные симметричные блоковые шифры как DES, S-DES, 3-DES, ГОСТ-89, ГОСТ-2015, AES, (См. [9]), очевидно, не являются КШ. Докажем далее, что таким свойством обладает несимметричный шифр Райвеста – Шамира – Адлемана (РША).

Шифр РША задается следующими параметрами:
p, q – различные простые числа (генерируются случайно);
$n = p * q$ – модуль алгоритма;
$\varphi = (p-1)(q-1)$ – функция Эйлера.

Ключ шифрования e генерируется случайным образом, но он должен удовлетворять условию: $1 \le e < \varphi, gcd(e, \varphi) = 1$, где $gcd(a, b)$ означает *наибольший общий делитель* чисел a и b; ключ дешифрования $d = e^{-1} \bmod n$.

Сообщение M, подвергающееся шифрованию, должно быть целым положительным числом в диапазоне: $1 \le M \le n - 1$. Процедура шифрования РША имеет вид: $E = M^e \bmod n$.

Тогда условие (10) принадлежности РША к классу СЕ будет иметь вид:

$$(M^{e1} \bmod n)^{e2} \bmod n = (M^{e2} \bmod n)^{e1} \bmod n. \qquad (11)$$

Очевидно, предполагается, что все модули в (11) совпадают.

Используя в (11) коммутативность возведения целых чисел в степень по любому модулю и коммутативность их произведения, получим в левой части (11):

$$(M^{e1} \bmod n)^{e2} \bmod n = M_3^{e1*e2} \bmod n = M^{e2*e1} \bmod n.$$

Для правой части (11) аналогично будем иметь:

$$(M^{e2} \bmod n)^{e1} \bmod n = M^{e2*e1} \bmod n.$$

Видно, что левая и правая части (11) одинаковы, и, следовательно, шифр РША принадлежит к классу КШ. Проиллюстрируем выполнение условия (11) для шифра РША простейшим примером. Пусть параметры РША заданы как: $p = 3$, $q = 5$, $n = 15$, $M = 5$, $e_1 = 3$, $e_2 = 5$. Тогда легко убедиться, используя обычные операции по модулю, что обе части (11) оказываются равными 8, и, следовательно, условие (10) выполняется.

Рассмотрим теперь пример другой несимметричной криптосистемы – Рабина [9]. Она имеет следующие параметры: p, q – различные простые числа, которые играют роль секретного ключа дешифрования $n = p * q$ (открытый ключ шифрования). Сообщение M представляется целым положительным числом в диапазоне $0 \le M \le n - 1$. Алгоритм шифрования имеет вид: $E = M^2 \bmod n$. Данная криптосистема имеет следующую особенность по сравнению с РША – стойкость взлома этой криптосистемы строго эквивалентна задаче факторизации целых чисел, т. е. нахождению множителей p, q при заданном n; такие криптосистемы принято называть *доказуемо стойкими*. Таким образом, если найден алгоритм факторизации целых чисел с полиномиальной сложностью от длины ключа, это означает, что найден алгоритм взлома (нахождение секретного ключа) криптосистемы Рабина, и, наоборот (что особенно важно!) – если удается как-то взломать криптосистему Рабина с полиномиальной сложностью от длины ключа, то это означает, что найдется и алгоритм факторизации целых чисел с полиномиальной сложностью от их разрядности. (Заметим, что таким свойством не обладает криптосистема РША). Однако, к сожалению, шифр Рабина не принадлежит к классу КШ, что доказывается следующим простым примером.

Свойство СЕ для данной криптосистемы можно записать следующим образом:

$$(M^2 \bmod n_1)^2 \bmod n_2 = (M^2 \bmod n_2)^2 \bmod n_1 \quad (12)$$

Выберем следующие параметры криптосистемы Рабина: $p_1 = 3$, $q_1 = 5$, $n_1 = 15$, $p_2 = 11$, $q_2 = 7$, $n_2 = 77$, $M = 5$. Подставляя эти параметры в (12), получим, что левая часть (12) равна 23, а правая – 10. Это доказывает, что шифр Рабина не принадлежит к КШ.

В первой части настоящей статьи отмечалось, что появление квантовых компьютеров (КК) создает опасность как для симметричных, так и для несимметричных криптосистем, и, в частности, для шифров РША, стойкость которых зависела от возможности полиномиально-сложного решения задачи дискретного логарифмирования или факторизации чисел. Поэтому появился целый класс так называемых *постквантовых* криптосистем, т. е. таких, которые не могут быть взломаны даже при практическом появлении КК. Первым представителем такого класса явилась криптосистема Мак-Элиса [9].

Покажем далее, что криптосистема Мак-Элиса, к сожалению, не использует шифры, относящиеся к классу КШ.

Процедура шифрования для криптосистемы Мак-Элиса описывается следующим образом:

$$E = M\tilde{G} \oplus Z,$$

где M – столбец сообщений длиной «K»; $\tilde{G} = SGP$ – матрица размерности $K \times n$, которая является открытым ключом шифра Мак-Элиса; S – случайная, несингулярная $K \times K$ матрица; G – случайная порождающая матрица Гоппа кода [9]; P – случайная $n \times n$ перестановочная матрица; Z – случайный двоичный вектор длины n и заданного веса t.

Матрицы S, G, P являются секретным ключом дешифрования криптосистемы Мак-Элиса.

Заметим, что алгоритм шифрования этой криптосистемы является *рандомизационным*, поскольку вектор Z выбирается случайно и не входит в состав ключей криптосистемы Мак-Элиса.

Для того, чтобы шифр Мак-Элиса был бы коммутативным, он должен удовлетворять условию (10), которое для данной криптосистемы принимает следующий вид:

$$f_{k1}(f_{k2}(M)) = \left(M\tilde{G}_2 \oplus Z_2\right)\tilde{G}_1 \oplus Z_1 = \\ = M\tilde{G}_2\tilde{G}_1 \oplus Z_2\tilde{G}_1 \oplus Z_1, \quad (13)$$

$$f_{k2}(f_{k1}(M)) = \left(M\tilde{G}_1 \oplus Z_1\right)\tilde{G}_2 \oplus Z_2 = \\ M\tilde{G}_1\tilde{G}_2 \oplus Z_1\tilde{G}_2 \oplus Z_2. \quad (14)$$

Легко видеть, что правые числа в (13) и (14) не совпадают, поскольку $Z_1 \neq Z_2$. Ввиду этого шифр Мак-Элиса не принадлежит к классу коммутативных криптосистем.

Рассмотрим следующее несимметрическое шифрование на предмет его принадлежности к КШ – это шифр, построенный на *цифровых решетках*. Как известно [24], один из популярных вариантов такого шифра, с аббревиатурой LWE (*аббр. от англ.* Learning With Errors; *перев. на русск.* обучение с ошибками), имеет следующий алгоритм шифрования:

$$\bar{u} = A^{\mathrm{T}}\bar{a},\ \bar{c} = P^{\mathrm{T}}a + f(v) \in Z_q^n x Z_q^l,$$

т. е. криптограмма такой криптосистемы состоит из двух цепочек с целочисленными координатами общей длиной $n + l$, тогда как сообщение $v \in Z_q^n$ – таким образом, вообще невозможно с ее помощью производить повторное шифрование. Поэтому шифр LWE не принадлежит к классу КШ.

Сравнительно недавно был предложен такой новый класс криптосистем с открытым ключом, как *некоммутативная криптография* [23]. Однако решено было не менять заголовок настоящего раздела статьи, несмотря на его очевидную формальную схожесть с упомянутым выше термином, тем более, что у них имеется и очевидное различие. Действительно, КШ относится лишь к возможности перестановки ключей при двойном шифровании, тогда как термин «некоммутативная криптография» относится ко всей структуре криптосистем, базирующейся на некоммутативности определенных неабелевых групп, использующихся при построении всей криптосистемы.

Интересно отметить, что потоковые шифры являются коммутативными! Действительно, условие (10) для них тривиально выполняется:

$$M \oplus \gamma(K_1)) \oplus \gamma(K_2) = (M \oplus \gamma(K_2)) \oplus \gamma(K_1),$$

где $\oplus$ – операция побитового сложения по mod2; $\gamma(k)$ – двоичные гаммы, как функции секретных ключей.

Протокол передачи секретной информации без предварительного обмена секретными ключами, представленный на рисунке 3, для потоковых шифров имеет вид, показанный на рисунке 4.

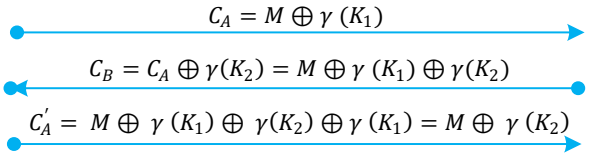


Рис. 4. Протокол, использующий КШ для потоковых шифров

Fig. 4. Protocol Using CE with Stream Ciphers

Однако из схемы (см. рисунок 4) видно, что перехватчик, побитно складывая по mod2 двоичные последовательности C_B и C_A', получает:

$$C_B \oplus C'_A = M \oplus \gamma(K_1) \oplus \gamma(K_2) \oplus M \oplus \gamma(K_2) = \gamma(K_1). \quad (15)$$

Затем перехватчик, складывая отдельно принятую им последовательность C_A и последовательность, которая была им вычислена по (15), дешифрует сообщение M:

$$M \oplus \gamma(K_1) \oplus \gamma(K_1) = M,$$

без всякого взлома гаммы потокового шифра...

Таким образом, для потокового шифра, даже при его коммутативности, протокол, представленный на рисунке 3, не работает, что объясняется линейностью для него процедуры шифрования.

Из всего вышеизложенного, можно сделать важное заключение, которое составляет одну приведенную далее проблему криптографии – доказать (или опровергнуть) утверждение, что существует постквантовый коммутативный шифр.

Поясним, однако, важно ли, вообще, заниматься решением этой проблемы? Это будет следовать из понимания того, имеет ли протокол, показанный на рисунке 3, существенное преимущество на практике. Для этого необходимо уточнить положительные и отрицательные свойства данного протокола. Поскольку пока не найдено ни одной симметричной криптосистемы, которая была бы также и коммутативной, т. е. удовлетворяла бы условию (10), то рассмотрим особенности использования КШ для несимметричных шифров.

Основное преимущество КШ состоит в том, что в этом случае не требуется никакой предварительной передачи ключей по каналам связи, в частности и открытых ключей, как, например, для КШ РША. Это может восприниматься, как дополнительный фактор повышения секретности. Важно подчеркнуть, что как КШ, так и не КШ, требуют защиты от *атак имперсонализации* активного злоумышленника, что, в свою очередь, нуждается в использовании методов аутентификации. Действительно, если такой злоумышленник выдаст себя за легитимного пользователя B и выполнит под видом B второй шаг протокола, показанного на рисунке 4, то ничего не подозревающий пользователь A , выполнит третий шаг протокола, предоставив злоумышленнику возможность просто расшифровать сообщение M , предназначенное A для легитимного пользователя B. Однако для КШ требуется *идентификация пользователей*, а для обычных несимметричных криптосистем шифрования сообщений требуется *аутентификация открытых ключей.* В этом и состоит преимущество КШ перед криптосистемами шифрования с открытыми ключами.(Собственно говоря, здесь атака имперсонизации ничем не отличается от подобной атаки при выполнении протокола распределения ключей Диффи – Хеллмана [9]).

4. Протокол Дина и Голдсмит передачи секретной информации по каналам с замираниями без предварительного обмена ключами

Криптосистема, предложенная Дином и Голдсмит [25], также принадлежит к классу *физически секретного уровня* (см. в [12] – «Physual layer security»), поскольку секретность передаваемых сообщений обеспечивается для нее естественными физическими свойствами канала связи без всякого использования шифрования / дешифрования и распределения ключей. Однако она может быть реализована не для любого канала связи, причем даже и с естественными шумами, но лишь для каналов связи с замираниями и с использованием технологии MIMO (*аббр. от англ.* Multiple-Input, Multiple-Output; *досл. перев. на русск.* множественный вход, множественный выход, по сути – это метод пространственного кодирования сигнала). Реально это соответствует использованию множества управляемых антенн на передаче и на приеме, а также при размещении легитимных пользователей и перехватчика в различных (и не совпадающих) точках пространства, что приводит к различным статистическим характеристикам каналов передачи, подвергающимся случайным изменениям.

Предположим сначала, что количество приемных антенн легитимных пользователей n_r и перехватчика n'_r одинаково (впоследствии это ограничение будет снято).

Пусть легитимный канал связи от корреспондента A к корреспонденту B описывается следующим выражением:

$$z = Ay + E, \quad (16)$$

где $A \in R^{n\times n}$ – матрица, описывающая канал связи $A \to B$; $z \in R^n$ – вектор сигнала, принимаемого B; $y \in R^n$ – вектор сигнала, переданного A; $E \in R^n$ – вектор аддитивного шума у B.

Предполагается также, что E – это гауссовские $i.i.d.$ векторы с параметрами $N(O, \sigma_E^2)$; элементы матрицы $A - (a_{ij})$, $i = (1, n)$, $j = (1, n)$ $i.i.d.$ – гауссовские $N(O, \sigma^2)$ величины.

Аналогичный (по структуре) канал перехвата $A \to E$ описывается выражением:

$$z' = By + E', \quad (17)$$

которое отличается от выражения (16) A лишь матрицей B и параметрами $\tilde{\sigma}_\varepsilon^2$ и $\tilde{\sigma}_w^2$.

Принятые для моделей каналов условия являются весьма важными и состоят в следующем:

– параметры каналов постоянны в течение выполнения процедур кодирования и декодирования;

– матрица A известна в точности легитимным пользователям;

– матрица A и B в точности известны перехватчику E.

В соответствии с [25] кодирование сообщения пользователем A выполняется следующим образом:

$$Y = Vx, \tag{18}$$

где $V \in R^{n\times n}$ – ортогональная матрица, взятая из SVD (*аббр. от англ.* Singular Value Decomposition; *перев. на русск.* сингулярное разложение), $A = USV^T$. Здесь $x \in R^n$ – вектор с двоичными координатами.

Декодирование легитимным пользователем B выполняется в два этапа. Первый этап – преддекодирование:

$$\begin{aligned} z' &= U^T z = U^T A y \\ +e' &= U^T U S V^T V x + U^T e = Sx + e', \end{aligned} \tag{19}$$

где $U \in R^{n\times n}$ – ортогональная матрица, взятая B из SVD-разложения матрицы A.

Поскольку S – диагональная матрица (по определению SVD), то на втором этапе декодирование B выполняет процедуру:

$$x' = \underset{x_i}{\operatorname{argmin}} |z_i - x_i\, S_i|,\ i = (1, n). \tag{20}$$

Из выражения (20) видно, что декодирование для легитимного пользователя B имеет линейную сложность в зависимости от параметра n.

По предположению авторов статьи [25], перехватчик E может повторить стратегию декодирования легитимного пользователя B, т. е.:

$$z'' = U^T Z'. \tag{21}$$

Тогда

$$z'' = U^T B y + E'' = U^T S' V'^T V x + e''. \tag{22}$$

где U', V', S' – матрица из SVD-разложения матрицы B.

Наконец, запишем (22) следующим образом:

$$z'' = Cx + e'', \tag{23}$$

где $C = S'V'^T V$.

Однако, поскольку матрица C не диагональна, то в этом случае оптимальное декодирование принимает вид:

$$x' = \underset{x_i}{\operatorname{argmin}} \|z'' - Cx_i\|, \tag{24}$$

где $||*||$ – евклидова норма в R^n.

Решение задачи (24) известно, как трудная CVP-проблема, и в [25] доказано, что она имеет экспоненциальную сложность от параметра n, если выполняется следующее (довольно очевидное) условие:

$$\sigma_w^2 * \tilde{\sigma}_e^2 > n^{\frac{1}{2}}.$$

Поэтому авторы статьи [25] сделали вывод, что предложенная ими криптосистема является стойкой, по крайней мере, если элементы матриц A и B являются взаимно независимыми случайными величинами. Это же условие обеспечивается, в свою очередь, при выполнении другого условия, когда расстояние от расположения легитимных пользователей A и B до перехватчика оказывается не меньше, чем несколько длин волн радиоканала $A \to B$.

Однако авторы статьи [25] не учли того обстоятельства, что перехватчик не обязательно должен следовать протоколу декодирования легитимных пользователей.

В работе [21] было предложено использовать следующий субоптимальный алгоритм декодирования при довольно вероятном условии, что матрица C из (24) окажется несингулярной:

$$C^{-1} z'' = x + C^{-1}\tilde{e}. \tag{25}$$

Из выражения (25) видно, что тогда процедура декодирования перехватчиком E принимает следующий вид:

$$x' = \underset{x_i}{\operatorname{argmin}} |\tilde{z}_i - x_i|, i = (z, n). \tag{26}$$

где $\tilde{z}_i$ – это i-я координата вектора $C^{-1}z''$.

Очевидно, что алгоритм декодирования по (26) имеет линейную сложность от параметра n, и тогда эффективность субоптимального декодирования можно будет оценить сравнением вероятности ошибки бита при оптимальном декодировании по (24) и субоптимального декодирования по (26).

В таблице 2 представлены результаты расчетов полученные в [21] для вероятностей ошибки P при оптимальном декодировании по (24) и при субоптимальном декодировании P' по (26) при «типичных» параметрах каналов связи $A \to B$ и $A \to E$.

ТАБЛИЦА 2. Результаты моделирования P и P' для типичных наборов параметров каналов $A \to B$ и $A \to E$

TABLE 2. Simulation Results of P and P 'for Typical Channel Parameters A → B and A → E

Наборы параметров	Для легитимных пользователей (P)	Для перехватчика (P')
$\sigma^2 = \sigma_w^2 = 2;\ \sigma_E^2 = \tilde{\sigma}_e^2 = 1;\ n = 100$	0,02	0,2
$\sigma^2 = \sigma_w^2 = 4;\ \sigma_E^2 = \tilde{\sigma}_e^2 = 8;\ n = 100$	0,037	0,3
$\sigma^2 = \sigma_w^2 = 1;\ \sigma_E^2 = \tilde{\sigma}_e^2 = 30;\ n = 100$	0,02	0,42
$\sigma^2 = 2; \sigma_w^2 = 1; \sigma_E^2 > \tilde{\sigma}_e^2 = 4;\ n = 1000$	$5{,}6 \cdot 10^{-3}$	0,3
$\sigma^2 = 4; \sigma_w^2 = 8; \sigma_E^2 = \tilde{\sigma}_e^2 = 12;\ n = 1000$	0,01	0,33

Таблица 2 показывает, что оптимальный алгоритм (24) дает вполне приемлемые результаты по вероятности ошибки символа, в то время как подоптимальный – приводит к недопустимо большим вероятностям ошибок. Более того, в работе

[21] строго доказано, что, даже используя избыточность сообщений (например, смысловой текст), его окажется невозможно восстановить с высокой надежностью.

Однако возникает очевидный вопрос – сохранится ли данное утверждение, если количество приемных антенн n_r' в MIMO-системе связи для перехватчика оказывается больше, чем количество приемных антенн n_r легитимной линии связи? Такое исследование было проведено в работе [21], и его результаты приведены в таблице 3, причем матрица C, получаемая при подоптимальном декодировании, теперь уже оказывается не квадратной, а прямоугольной, и поэтому для ее обращения (т. е. нахождения обратной матрицы C^{-1}) используется известный *алгоритм Мура – Пенроуза* [26].

ТАБЛИЦА 3. Результаты моделирования оптимального и подоптимального алгоритмов декодирования для параметров $n_r' > n_r$, $n_r = 100$

TABLE. 3 .Simulation Results of Optimal and Suboptimal Decoding Algorithms with Parameters: $n_r' > n_r$, $n_r = 100$

n_r'	100	101	102	103	105	107	108	109	110	120	150
P'	0,31	0,22	0,16	0,12	0,07	0,048	6,039	0,03	0,024	0,003	$7 \cdot 10^{-4}$

По результатам таблицы 3 можно сделать неожиданный вывод, что даже небольшое увеличение числа приемных антенн перехватчика создает возможность «взлома» данной криптосистемы.

В работе [27] было предложено изменить матрицу предкодирования V на матрицу A^{-1}, обратную матрице канала $A \to B$. Однако это приводит к требованию значительного увеличения мощности передаваемого сигнала для пользователя A (см. таблицу 5 в работе [21]). Кроме того, как отмечено в той же работе, неточность измерения матриц A и B легитимными пользователями приводит к значительному увеличению вероятностей ошибок в основном канале (см. таблицу 6 в работе [21]).

В заключение описания метода Дина и Голдсмит, можно высказать мнение, что, хотя дальнейшее теоретическое изучение этого метода представляет определенный интерес, особенно в направлении улучшения алгоритма преддекодирования, однако, с точки зрения практического применения, он представляется сомнительным. Тем более, что не всегда легитимным пользователям удается тщательно следить за изменениями характеристик канала связи $A \to E$.

5. Обеспечение секретности передачи ключевых данных в многолучевых каналах связи за счет использования антенн с управляемыми диаграммами направленности

Результаты по данному направлению были впервые получены японскими авторами и опубликованы в статье [28]. Значительное уточнение основных результатов и условий обеспечения секретности передачи ключевой информации (в терминах шенноновского количества информации) представлено в работе [29].

В настоящей статье будет описана модель, использованная для передачи ключевой информации, и сформулированы возможные направления исследований, поскольку, насколько нам известно, несмотря на многочисленные публикации по теории антенн с управляемой диаграммой направленности (VDA, *аббр. от англ.* Variable Directional Antenna), данная технология не развивалась российскими исследователями для обеспечения секретности передачи ключевых данных, а только лишь для повышения надежности радиосвязи в многолучевых каналах.

Схема передачи ключевых данных в условиях возможного их перехвата, на основе технологии VDA, приведена на рисунке 5.

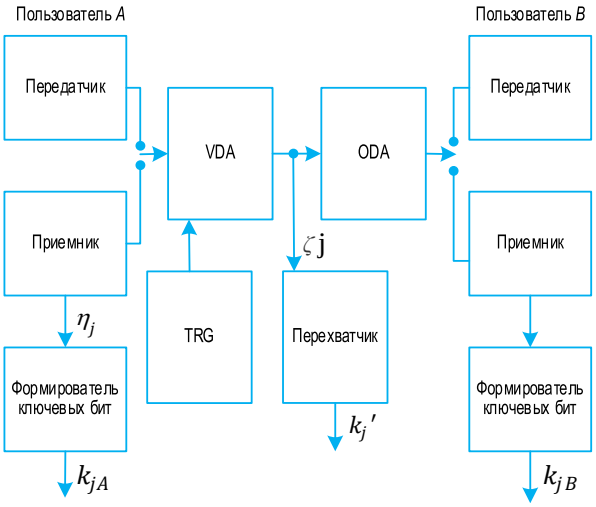


Рис. 5. Схема системы связи для передачи ключевых данных при использовании технологии VDA

Fig. 5. Block Scheme of Communication System for Key Transmission under the Use of VDA Technology

Протокол распределения ключевых данных описывается следующими шагами.

<u>*Шаг 1*</u>. Пользователь A формирует управляемую диаграмму направленности (VDA), используя чисто случайный генератор (TRG, *аббр. от англ.* Truly Random Generator) и фиксирует (запоминает) ее в течение временно́го интервала передачи (T) j-го ключевого бита.

<u>*Шаг 2*</u>. A передает B гармонический сигнал $S_j(\mathrm{t}) = \cos\omega_0 t$, $0 \le t \le T/2$ по многолучевому каналу.

<u>*Шаг 3*</u>. B принимает гармонический сигнал от всенаправленной антенны (ODA, *аббр. от англ.* Omni Directional Antenna) на интервале $(0, {}^{T}/_{2})$ и формирует ключевой бит, вычисляя выбранный функционал (амплитуду или фазу) сигнала.

Шаг 4. *B* также переключает свою ODA в режим передачи и передает тот же самый гармонический сигнал на интервале $T/2 \le t \le T$.

Шаг 5. Пользователь *A* подключает свою VDA (с сохраненной диаграммой направленности антенны) к приемнику и обрабатывает полученный сигнал также, как и *B*, выделяя *j*-й ключевой бит.

Шаг 6. *A* и *B* повторяют шаги 1–5 *n* раз с различными выходами генератора, чтобы создать желаемое число разделенных бит ключа между *A* и *B*.

Вследствие справедливости теоремы обратимости распространения радиоволн, ключевые последовательности пользователей *A* и *B* должны совпадать с точностью до шумов их приемников.

Поэтому сигналы, разделенные *A* и *B*, а также *A* и *E* будут иметь вид, соответственно:

$$y_j(t) = \sum_{i=1}^{m} v_{ij}\beta_{ij}\cos(\omega_{0t} + \theta_{ij}), \qquad (27)$$

$$z_j(t) = \sum_{i=1}^{m} v'_{ij}\,\beta'_{ij}\cos(\omega_{0t} + \theta'_{ij}), \qquad (28)$$

где *i* – *i*-й луч канала связи; *m* – общее количество лучей канала связи; β_{ij} – коэффициенты усиления каналов; v_{ij} – коэффициенты усиления VDA; θ_{ij} – фазовый сдвиг, включающий как фазу в антенной диаграмме, так и фазовый сдвиг в *i*-ом луче; штрихи над символами относятся к каналу перехвата.

В работе [29] было показано, что величины η_j и ζ_i могут быть достаточно хорошо аппроксимированы гауссовскими, а вероятность ошибки P_e между легитимными пользователями *A* и *B* и перехватчиком *E*, для функционалов выделения фаз сигналов, может быть представлена следующей формулой:

$$P_e = \frac{1}{\pi}\arctan\left(\frac{\sqrt{1-\rho^2}}{\rho}\right), \qquad (29)$$

где ρ – коэффициент корреляции между случайными величинами η_j и ζ_j.

На рисунке 6 представлена зависимость *Pe* от ρ. Интересно отметить, что вопреки нашей интуиции, вероятность расхождения бит ключа для легитимных пользователей и перехватчика, равная 0,1, может достигаться даже при коэффициенте корреляции $\rho \approx 0{,}95$!

Для существенного уменьшения утечки информации к перехватчику может быть использована процедура усиления секретности, подробно описанная в Разделе 2. Напомним, что формула для вычисления величины утечки информации к перехватчику в этом случае (т. е. для бесшумного основного канала) имеет вид:

$$I_0 \le \frac{-(k - t_c - t - r)}{\gamma_1 \ln 2}.$$

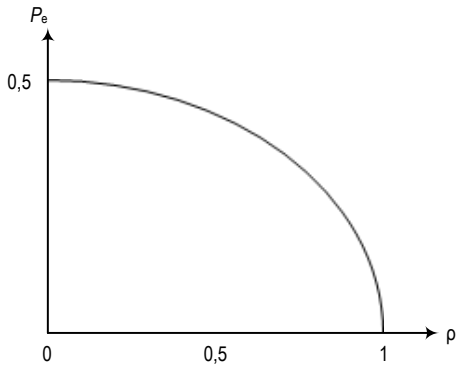


Рис. 6. Зависимость вероятности ошибки P_e от коэффициента корреляции ρ

Fig. 6. The Error Probability P_e versus Correlation Coefficient ρ

Параметры, входящие в формулу для I_0, описаны в формуле (8), где P_w, однако, вычисляется здесь по формуле (29).

При помощи теоретических расчетов и имитационного моделирования в работе [29] было показано, что метод, использующий VDA для распределения ключевых данных в многолучевом канале, при оптимизации параметров, позволяет обеспечить достаточно надежное распределение ключевых бит при малой утечке информации (по Шеннону) к перехватчику. Причем даже если канал перехвата является бесшумным, однако три следующих условия оказываются совершенно необходимыми для этого:

– хорошие свойства случайного генератора шума;

– наличие не менее двух лучей в канале с замираниями, как для основного, так и для канала перехвата;

– достаточная удаленность легитимных пользователей *A* и *B* от перехватчика *E*.

Заметим, однако, что эффективность рассмотренного метода зависит не только от расстояния между легитимными пользователями, но и от точного расположения последнего (*E*) относительно них.

6. Взлом криптосистемы EVESkey

В отличие от предыдущих, данный раздел поясняет необходимость тщательного анализа бесключевых систем современной криптографии с точки зрения их безопасности и невозможности присутствия различных побочных каналов, позволяющих обойти утверждение авторов об идеальной секретности предложенных ими бесключевых криптосистем. Типичным примером с таким пессимистичным сценарием является так называемая EVESkey-схема, предложенная в работе [30], которая характеризуется ее авторами, как абсолютно секретная. На рисунке 7 представлен сценарий, соответствующий данной схеме.

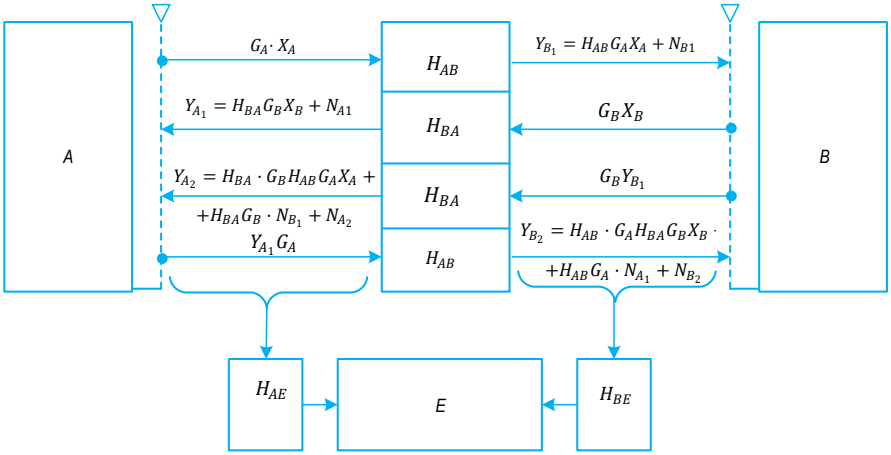


Fig. 7. Сценарий протокола для EVESkey-схемы

Fig. 7. Scenario of the Protocol for EVESkey-Scheme

Как видно из этой схемы, A и B генерируют случайные унитарные матрицы $X_A, X_B \in C^{n\times n}$, где n – количество антенн, использующихся этими пользователями), а также матрицы Дженибра [26] $G_A, G_B \in C^{n\times n}$. Матрицы H_{AB} и H_{BA} являются $n \times n$ передаточными матрицами каналов $A \to B$, $B \to A$, соответственно, а их элементы моделируются как взаимно независимые гауссовские величины $N(0,1)$.

Матрицы N_A, N_B положим $n \times n$ – матрицами аддитивного шума с элементами $N(0, \sigma^2)$.

Для упрощения доказательства введем вспомогательные матрицы $P = H_{BA}G_B$ и $Q = H_{AB}\,G_A$.

Тогда PQ и QP могут быть оценены легитимными пользователями как:

$$PQ = Y_{A2}(X_A)^{-1};\ QP = Y_{B2}(X_B)^{-1}. \qquad (30)$$

Из алгебры хорошо известно, что любые матрицы PQ и QP имеют одинаковые характеристические полиномы (CP, *аббр. от англ.* Characteristic Polynomials), т. е.:

$$CP(PQ) = CP(QP). \qquad (31)$$

Из равенства (31), очевидно, следует, что легитимные пользователи A и B, получая оценки матриц Y_{A2} и Y_{B2} и имея в своем распоряжении матрицы X_A и X_B, могут легко восстанавливать матрицы PQ и QP, соответственно, а затем, вычислив из CP собственные числа $EV(PQ)$ и $EV(QP)$ этих матриц и проквантовав их, получить с большой вероятностью совпадающие ключевые биты. В то же время перехватчик E не обладает знанием матриц X_A или X_B и поэтому не может воспользоваться одним из равенств (30), чтобы затем найти ключевые биты и после проквантовать $EV\,(QP)$ и $EV\,(PQ)$.

Покажем, однако, что протокол вычисления, описанный выше, может быть взломан, если не следовать строго протоколу, предписанному легитимным пользователям, а воспользоваться лишь теми данными, которые перехватчик сможет получить, наблюдая его выполнение легитимными пользователями, правда в предположении, что аддитивные шумы легитимных пользователей A, B и перехватчика E отсутствуют.

Тогда E (см. рисунок 7) сможет вычислить следующие матрицы:

$$\tilde{Y}_{A1} = H_{BE}G_BX_B, \tilde{Y}_{A2} = H_{BE}G_BH_{AB}G_AX_A,$$
$$\tilde{Y}_{B1} = H_{AE}G_AX_A, \tilde{Y}_{B2} = H_{AE}G_AH_{BA}G_BX_B. \qquad (32)$$

Докажем далее два утверждения.

Утверждение 1

Для всех стоящих в левых частях выражения (32) случайных матриц (в общем случае и для прямоугольных матриц) существует псевдообратная матрица Мура – Пенроуза $(\gamma_p)^{-1}$ с вероятностью единица.

Доказательство

Действительно, положим H – $n \times n$ матрица с комплексными элементами $N(0,1)$, которые взаимно независимы.

Тогда совместная плотность вероятности такой матрицы будет иметь вид:

$$f(H) = (2\pi)^{-n^2} \exp\left(-\frac{1}{2}Tr(H \times H^T)\right), \qquad (33)$$

где Tr (.) – след матрицы.

Вырожденные матрицы, у которых $detH = 0$ образуют $2n^2 - 2$ размерное многообразие, которое будучи несингулярным, обеспечивает нулевую вероятность для $detH = 0$. Таким образом, вероятность того, что $detH \neq 0$, будет равна единице. Подобные матрицы остаются обратимыми и будучи умноженными на унитарные матрицы G, X.

Утверждение 2

Обозначим через $EV(Y)$ множество собственных чисел матрицы Y. Тогда:

$$EV(Y) = EV(PQ) = EV(QP), \quad (34)$$

где

$$Y = \tilde{Y}_{A2}(\tilde{Y}_{B1})^{-1}\tilde{Y}_{B2}(\tilde{Y}_{A1})^{-1}. \quad (35)$$

Доказательство

Подставляя (32) в (35), получим:

$$Y = (H_{BE}G_B)QP(H_{BE}G_B)^{-1}.$$

Последнее выражение означает, что матрица Y подобна матрице QP, и тогда $EV(Y) = EV(QP)$ для любых матриц H_{BE}. Поэтому, вычислив $EV(Y)$, где Y находится по (35) и (32), перехватчик E находит ключевые биты, распределенные по данному протоколу A и B. Что и требовалось доказать.

Хотя ранее предполагалось, что все каналы передачи информации считаются бесшумными, но дополнительные расчеты показали, что и при не слишком больших мощностях шума вероятность битовых ключевых ошибок легитимных пользователей оказывается достаточно близкой к вероятности ошибки ключевых бит перехватчика E, выполняющего протокол, используя (34) и (35). Из этого следует, что схему распределения ключей EVESKey нельзя считать секретной. Интересно отметить, что авторы настоящего исследования посылали в журнал, где ранее опубликовали статью [30], материал, полностью компрометирующий EVESkey-схему, однако редакторы этого журнала, ссылаясь на дополнительные комментарии рецензентов, отказались опубликовать наше опровержение.

Подводя итоги различным методам бесключевой криптографии, описанным в разделах 2–6 настоящей работы, можно заметить, что все они требуют выполнения некоторых условий, относящихся к перехватчику.

Для удобства пояснения, все они сведены в таблицу 4, их которой видно, что требования зависят от возможностей перехватчика и сведений о его канале, что не всегда позволяет их выполнить. (Тем более, что они могут иногда скрытно изменяться). Поэтому возникает проблема – разработать протокол распределения ключей по постоянным, открытым и бесшумным каналам связи (типа Интернет). Попытка ее решения была предпринята в нескольких работах авторов настоящей статьи, и к обсуждению этих работ мы сейчас и переходим.

ТАБЛИЦА 4. Условия выполнимости секретных протоколов передачи данных по открытым каналам связи между легитимными пользователями методами бесключевой криптографии

TABLE 4. Conditions of Secret Data Transmission Availability over Communication Channels between Legitimate Users by the Means of Keyless Cryptography

Вайнеровская концепция отводного канала связи	Коммутативное шифрование	Протокол Дина и Голдсмит	Использование антенн с управляемой диаграммой
Необходимо знание, по крайней мере, мощности шума в канале перехвата, которая не должна быть менее определенного порога	Пока известна лишь одна такая криптосистема – РША, которая не является, однако, постквантовой	Необходимо выполнение условия, что количество антенн перехватчика в сценарии MIMO не превосходит количества антенн легитимных пользователей, использующих такой же сценарий MIMO	Необходимо использовать антенну с управляемой диаграммой направленности, а канал передачи должен иметь не менее 2-х лучей и случайное затухание сигналов

7. Протокол распределения крипто ключей, использующий постоянные, открытые и бесшумные каналы связи (типа Интернет)

В работе [31] был описан протокол распределения ключей, выполняемый по постоянному, открытому и бесшумному двоичному каналу связи при помощи передачи двоичных последовательностей. Схема такого протокола приведена на рисунке 8.

Как видно из рисунка 8, A и B предварительно генерируют чисто случайные двоичные последовательности δ_A, γ_A и δ_B, γ_B, соответственно, а затем обмениваются ими по каналам $A \to B, B \to A$ и генерируют предварительные ключи:

$$\tilde{K}_A = \delta_A \oplus \delta_B \oplus \gamma_B,\ \tilde{K}_B = \delta_B \oplus \delta_A \oplus \gamma_A,$$

где $\oplus$ – операция побитового сложения по mod2.

Перехватчик E вырабатывает подобный предварительный ключ:

$$\tilde{K}_E = \delta_A \oplus \delta_B \oplus \gamma_A \oplus \gamma_B.$$

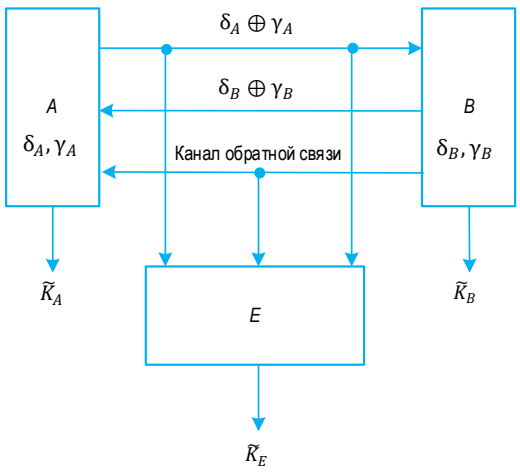


Рис. 8. Протокол распределения ключей с использованием обмена информации по двоичному бесшумному каналу

Fig. 8. Key Sharing Protocol with the Use of Information Exchange over Binary Noiseless Channel

Далее A и B используют для обеспечения хорошей статистики финального ключа γ дополнительную схему, приведенную на рисунке 9. Как видно из этой схемы, в качестве общего ключа между легитимными пользователями принимается последовательность γ, которая для B искажается шумом $\gamma_A \oplus \gamma_B$, а для E – шумом γ_B.

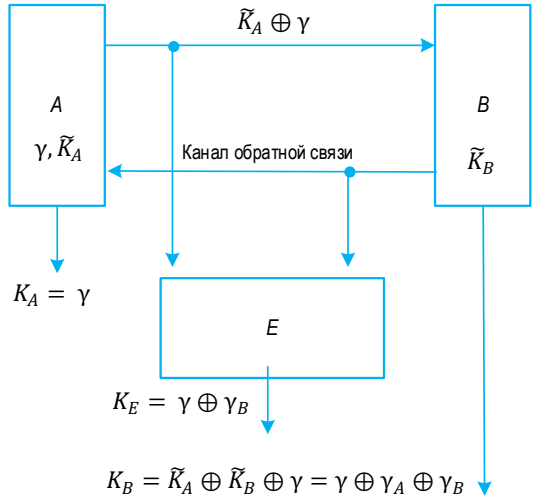


Рис. 9. Дополнительная схема для выработки финального ключа между *A* и *B*

Fig. 9. Additional Scheme for Elaboration of the Final Key between A and B

Схема на рисунке 9 эквивалентна схеме, приведенной на рисунке 10. В [32] схема, подобная приведенной на рисунке 10, названа *каналом с отводом при деградации основного канала.*

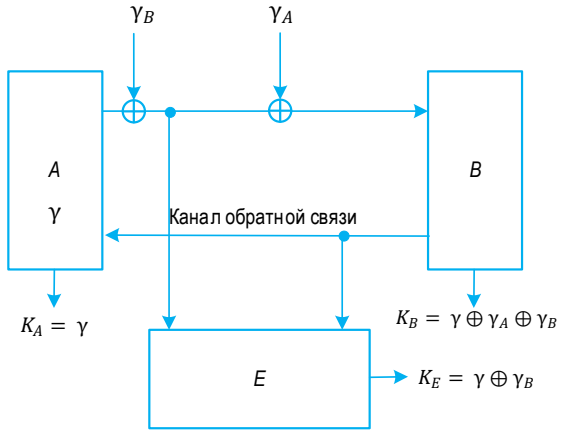


Рис. 10. Эквивалентная схема для схемы на рисунке 9

Fig. 10. Equivalent Scheme for the Scheme Shown in Fig,9

Более того, в той же работе [32] строго доказано, что для такой модели *секретная пропускная способность* $C_S = 0$. (Напомним, что по определению, данному в [14], секретной пропускной способностью, в терминах шенноновской информации, называется пропускная способность основного канала при обеспечении сколь угодно малой утечки по каналу с отводом). Это означает, что не существует никаких схем кодирования и декодирования, которые бы для данной модели обеспечили надежную передачу информации (в нашем случае ключевой) между легитимными пользователями при отсутствии утечки к перехватчику.

Рассмотрим другой протокол распределения ключей, который показан на рисунке 11, где обмен между легитимными пользователями осуществляется случайными вещественными числами.

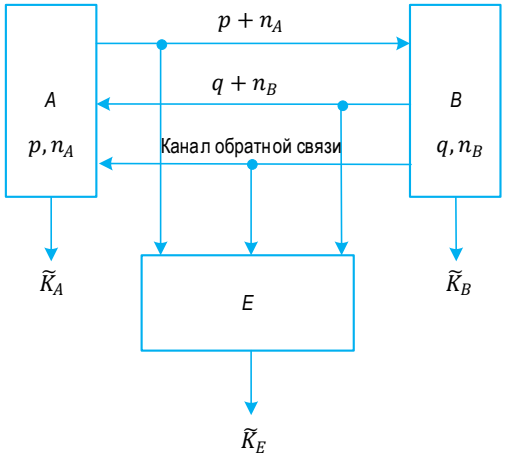


Рис. 11. Протокол распределения ключей при обмене по каналам связи вещественными числами

Fig. 11. Key Sharing Protocol for Communication Channel Exchange by Real Numbers

Предположим, что все случайные вещественные числа являются гауссовскими, взаимно независимыми, имеют нулевые матожидания, а дисперсии (вариации):

$$Var(p) = Var(q) = 1, Var(n_A) = Var(n_B) = \sigma^2.$$

После обмена по протоколу, показанному на рисунке 11, легитимные пользователи A и B формируют свои ключи по следующему правилу:

$$K_A = rect\big(p(q + n_B)\big),\ K_B = rect(q(p + n_A), \quad (36)$$

где «+» – обычное арифметическое сложение, а $rect(x) = \begin{cases} 0, \text{при } x \geq 0 \\ 1, \text{при } x < 0 \end{cases}$

Перехватчик E должен сформировать свой ключ, следуя выражению:

$$K_E = rect\big((p + n_A)(q + n_B)\big). \quad (37)$$

Используя легко доказываемое равенство:

$$rect(ab) = rect(a) \oplus rect(b),$$

где $\oplus$ – сложение по mod 2.

Схему, показанную на рисунке 11, можно заменить на схему, приведенную на рисунке 12, где все ключи вычисляются по правилам:

$$\begin{aligned} K_A &= rect(p) \oplus rect(q + n_B), \\ K_B &= rect(p + n_A) \oplus rect(q), \\ K_E &= rect(p + n_A) \oplus rect(q + n_B). \end{aligned} \quad (38)$$

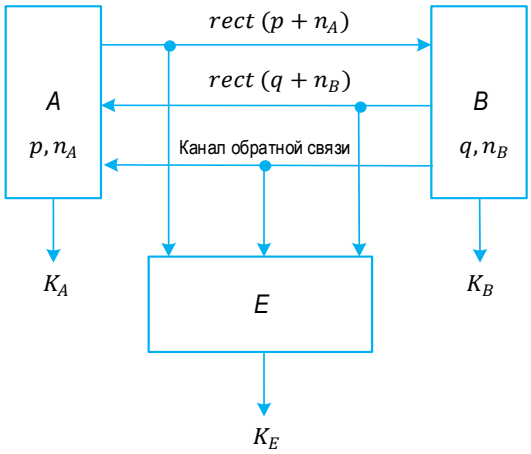


Рис. 12. Схема эквивалентная схеме на рисунке 11

Fig. 12. Scheme that is Equivalent to the Scheme in Fig.11

Найдем теперь аддитивные шумы между A и B (ε_{AB}) и между A и E (ε_{AE}):

$$\varepsilon_{AB} = K_A \oplus K_B = rect(p) \oplus rect(q + n_B) \oplus \oplus rect(p + n_A) \oplus rect(q),$$
$$\varepsilon_{AE} = K_A \oplus K_E = rect(p) \oplus rect(q) \oplus \oplus rect(p + n_A) \oplus rect(q + n_B) = = rect(p) \oplus rect(p + n_A). \quad (38)$$

Из (38) и (39) видно, что схема, приведенная на рисунке 12, будет эквивалентна схеме (см. рисунок 10), если в последней заменить γ_B и γ_A, соответственно, на:

$$rect(P) \oplus rect(P + n_A),\ rect(q) \oplus rect(q + n_B).$$

Но это означает, что схема на рисунке 12 согласуется со сценарием канала с отводом и деградацией основного канала. Тогда, как уже отмечалось при рассмотрении схемы на рисунке 8, в соответствии с результатами работы [32], мы получили и для схемы на рисунке 11 секретную пропускную способность $C_s = 0$, и, следовательно, выполнение по ней секретного протокола распределения ключей сказывается невозможным.

В работе [33] рассмотрен протокол при обмене между легитимными пользователями многомерными векторами с вещественными координатами. Однако мы не будем подробно рассматривать этот случай, поскольку для него не удалось строго доказать, что $C_s = 0$. Тем не менее, представляется, что это весьма вероятно, так как после применения этого протокола при моделировании не удалось получить необходимые условия для возможности применения теоремы усиления секретности, поскольку сохранилось прежнее неравенство: $P_m > P_w$, где P_m – вероятность битовой ошибки в основном канале; P_w – то же самое, но для канала перехвата.

Обратимся, наконец, к наиболее перспективному случаю, когда протокол выполняется при помощи обмена случайными матрицами, как это показано на рисунке 13. (Не умаляя общности, рассматриваем только квадратные матрицы, поскольку, по авторскому мнению, это не сможет существенно повлиять на фактор секретности протокола в терминах утечки к перехватчику определенного количества шенноновской информации).

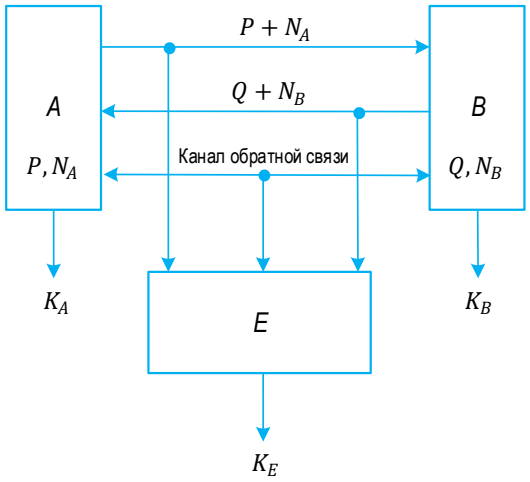


Рис. 13. KSP при обмене квадратными матрицами

Fig. 13. KSP under Square Matrices Exchange

Итак, пусть P, Q, N_A, N_B – $n \times n$ матрицы с гауссовскими взаимно независимыми элементами. Тогда, после выполнения протокола по схеме на рисунке 13 ключи легитимных пользователей A и B, а также перехватчика E вычисляются следующим образом:

$$K_A = rect(tr(P(Q + N_B))),$$
$$K_B = rect(tr(Q(P + N_A))),$$
$$K_E = rect\ (tr((P + N_A)(Q + N_b)),$$

где $tr(X)$ – след матрицы X (т. е. сумма ее диагональных элементов).

Заметим, что вместо функционала $tr(X)$ можно использовать и другие преобразования элементов матрицы. Так, в работе [31] были использованы $EV(X)$, т. е. собственные числа матриц (конечно, также с последующим их квантованием). В этой же работе было выполнено моделирование полного KSP и получены следующие обнадеживающие результаты: после применения LPDC-кодов и процедуры усиления секретности вероятность ошибочного приема хотя бы одного символа ключевого блока длиной 24039 бит оказалась равной $P_{ed} \approx \approx 2{,}5 \cdot 10^{-3}$, а утечка в этом блоке к перехватчику соответствовала $I_0 \approx 1{,}4 \cdot 10^{-3}$ бита. (Аналогичные по порядку результаты были получены и при выборе функционала tr вместо EV). Казалось бы, проблему распределения ключей по постоянному и бесшумному каналу можно считать полностью решенной, но это не так…

Дело в том, что до сих пор мы рассматривали обработку сигналов перехватчиком, как *жесткое декодирование*, т. е. при квантовании следов матриц

или их собственных чисел, на два уровня, которые сопоставляются двум битам – 0 и 1. Однако возможно и так называемое *мягкое декодирование* или квантование tr, EV или результатов применения других функционалов на $L > 2$ уровней! При этом перехватчик E может провести предобработку сигналов, т. е. составить специальную таблицу декодирования еще до получения результатов обмена данными между A и B. Пример такой таблицы представлен в работе [33]. Там же было произведено имитационное моделирование и всего соответствующего матричного протокола. Оно показало, что при оптимизации параметров квантования и декодирования перехватчик E получает преимущество перед легитимными пользователями, т. е. выполнение условия $P_w > P_m$, что не позволяет в дальнейшем, используя процедуру усиления секретности, обеспечить требуемую малую утечку информации при надежной передаче бит ключа от A к B. Конечно, составление таблицы предобработки требует значительно бóльшей сложности и / или бóльшего времени обработки, чем простое двоичное квантование. Поэтому подобный метод распределения ключей допустим лишь для легитимных пользователей не слишком высокого ранга, когда можно надеяться на жесткие ограничения сложности обработки сигнала перехватчиками. Более того, легитимные пользователи могут также заменить жесткое декодирование при обмене на мягкое, и тогда есть надежда, что в соревновании по сложности обработки, зависящем от числа уровней квантования, у них будет преимущество перед перехватчиком (см. также далее выводы, сформулированные в «Заключении»).

8. Заключение

В начале XX-го века Д. Гильберт сформулировал 23 математические проблемы, 16 из которых к настоящему времени уже решены в значительной степени, предопределив этим самым дальнейшее развитие математики. Заметим, что, в любой отрасли науки (физике, химии, биологии, а сейчас и в безопасности передачи цифровой информации), существуют свои, конечно, значительно меньшего масштаба проблемы, чем у «королевы наук» – математики. Однако их разрешение (или его невозможность его – см. труды К. Геделя) являются весьма важными для дальнейшего развития науки в данной отрасли – криптографии. Поэтому, отнюдь не претендуя по масштабу на «глобальность», по сравнению с проблемами Д. Гильберта, все же рискнем (по материалам обеих частей настоящей статьи) сформулировать весьма важные, по авторскому мнению, но, конечно, не единственные в области информационной безопасности, проблемы.

Во-первых, верхняя граница возможности использования одного и того же ключа шифрования без утечки информации к перехватчику для неидеального шифра (см. Часть 1).

Во-вторых, построение какой-либо постквантовой ассиметричной и коммутативной (в нашем смысле) криптосистемы (см. Раздел 3 Части 2);

В-третьих, вычисление секретной пропускной способности протокола распределения ключей при обмене информацией по постоянному, открытому и бесшумному каналу (см. Раздел 7 Части 2). Если окажется, что $C_S \neq 0$, – то остается проблема нахождения конструктивных методов кодирования и декодирования для такого сценария.

Список источников

1. Alpern B., Schneider F.B. Key exchange using ‘keyless cryptography’ // Information Processing Letters. 1983. Vol. 16. Iss. 2. PP. 79–81. DOI:10.1016/0020-0190(83)90029-7

2. Korzhik V. Keyless cryptography // Proceedings of the 9th International Conference on System Administration, Networking and Security (Orlando, USA). 2000.

3. Korzhik V., Bakin M. Information-theoretical Secure Keyless intensification // Proceedings of the 2000 IEEE International Symposium on Information Theory (Sorrento, Italy, 25–30 June 2000). Piscataway: IEEE Press, 2000. DOI:10.1109/ISIT.2000.866310

4. Korzhik V. Keyless cryptography // Invited Talk at Security Seminar at CERIAS Purdue University. 2001.

5. Mukherjee A., Fakoorian S.A.A., Huang J., Swindlehurst A.L. Principles of Physical Layer Security in Multiuser Wireless Network A. Survey // IEEE Communications Surveys & Tutorials. 2014. Vol. 16. Iss. 3. PP. 1550–1573. DOI:10.1109/SURV.2014.012314.00178

6. Wyner A.D. The Wire-tap channel // Bell System Technical Journal. 1975. Vol. 54. Iss. 8. PP. 1355–1387. DOI:10.1002/j.1538-7305.1975.tb02040.x

7. Bennett C.H., Bessette F., Brassard G., Salvail L., Smolin J. Experimental quantum cryptography // Journal of Cryptology. 1992. Vol. 5. PP. 3–28. DOI:10.1007/BF00191318

8. Кушнир Д.В. Исследование и разработка методов распределения конфиденциальных данных по квантовым каналам. Автореф. дис. канд. техн. наук. СПб.: СПбГУТ, 1996. 16 с. EDN:ZJDTTT

9. Коржик В.И., Яковлев В.А. Основы криптографии. СПб.: Интермедия, 2016. 296 с. EDN:WEQWMN

10. Korzhik V., Kushnir D. Key sharing based on the wire-tap channel type ii concept with noisy main channel // Proceedings of the International Conference on the Theory and Applications of Crypotology and Information Security (ASIACRYPT '96, Kyongju, Korea, 3–7 November 1996). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1996. Vol. 1163. PP. 210–217. DOI:10.1007/BFb0034848

11. Liu Y., Zhang W.J., Jiang C., Chen J.P., Zhang C., Pan W.X., et al. Experimental Twin-Field Quantum Key Distribution over 1000 km Fiber Distance // Physical Review Letters. 2023. Vol. 130. P. 210801. DOI:10.1103/PhysRevLett.130.210801

12. Milov M., Pham T.M., Chorti A., Barreto A.N., Fettweis G. Physical Layer Security – From Theory to Practice // IEEE BITS the Information Theory Magazine. 2023. Vol. 3. Iss. 2. PP. 67–79. DOI:10.1109/MBITS.2023.3338569

13. Шеннон К.Э. Работы по теории информации и кибернетике. М.: Издательство иностранной литературы, 1963. 829 с.

14. Maurer U.M. Secret key agreement by public discussion based on common information // IEEE Transactions on Information Theory. 1993. Vol. 39. Iss. 3. PP. 733–742. DOI:10.1109/18.256484

15. Коржик В.И., Яковлев В.А. Неасимптотическая оценка эффективности кодового зашумления в каналах с отводом // Проблемы передачи информации. 1991. № 4. С. 223–228.

16. Петерсон У., Уэлнод Э. Коды, исправленные ошибки. М.: Мир, 1976.

17. Коржик В.И., Яковлев В.А. Защита информации от утечки за счет побочных электромагнитных излучений и наводок на основе способа кодового зашумления // Информатика и вычислительная техника. 1993. Т. 8. № 1-2. С. 61–66.

18. Korzhik V., Morales-Luna G., Balakirsky V.B. Privacy amplification theorem for noisy main channel // Proceedings of the 4th International Conference on Information Security (ISC 2001, Malaga, Spain, 1–3 October 2001). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2001. Vol. 2200. PP. 18–26. DOI:10.1007/3-540-45439-X_2

19. Yakovlev V., Korzhik V., Morales Luna G. Key Distribution Protocol Based on Noisy Channels in Presence of Active Adversary // IEEE Transactions on Information Theory. 2008. Vol. 54. Iss. 6. PP. 2535–2550. DOI:10.1109/TIT.2008.921689

20. Шнайер Б. Прикладная криптография. М.: Триумф, 2002.

21. Korzhik V., Starostin V., Yakovlev V., Kabardov M., Krasov A., Adadurov S. Advance in Keyless Cryptography. Chapter 6 // In: Ramakrishnan S. (ed.) Lightweight Cryptographic Techniques and Cybersecurity Approaches. 2022. PP. 97–117. DOI:10.5772/intechopen.104429

22. Tilborg H.C.A. Encyclopedia of Cryptography and Security. Springer, 2005.

23. Myasnikov A.G., Shpilrain V., Ushakov A. Non-commutative Cryptography and Group-theoretic Problems. American Mathematical Society, 2011. 385 p.

24. Goldreich O., Goldwasser S., Halevi S. Public-key cryptosystems from Lattice reduction problems // Proceedings of the 17th Annual International Cryptology Conference (CRYPTO '97, Santa Barbara, USA, 17–21 August 1997). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1997. Vol. 1294. PP. 112–131. DOI:10.1007/BFb0052231

25. Dean T., Goldsmith Aj. Physical layer cryptography through massive MIMO // Proceedings of the 2013 IEEE Information Theory Workshop (ITW, 9–13 September 2013, Seville, Spain). Piscataway: IEEE, 2013. PP. 1–3. DOI:10.1109/ITW.2013.6691222

26. Ben-Israel A., Greville T.N.E. Generalized Inverses: Theory and Applications. Springer, 2003.

27. Steinfeld R., Sakzad A. On massive MIMO physical layer cryptosystems // Proceedings of IEEE Information Theory Workshop – Fall (ITW, Jeju, Korea (South), 11–15 October 2015). IEEE, 2015. PP. 292–296. DOI:10.1109/ITWF.2015.7360782

28. Aono T., Higuchi K., Ohira T., Komiyama B., Sasaoka H. Wireless secret key generation exploiting reactance-domain scalar response of multipath fadiny cannels // IEEE Transactions on Antennas and Propagation. 2005. Vol. 53. Iss. 11. PP. 3776–3784. DOI:10.1109/TAP.2005.858853

29. Korzhik V., Yakovlev V., Kovajkin Y. Secret Key Agreement Over Multipath Channels Exploiting a Variable-Directional Antenna // International Journal of Advanced Computer Science and Applications. 2012. Vol. 3. Iss. 1. PP. 172–178.

30. Qin D., Ding Z. Exploiting Multi Antenna Non-Reciprocal Channels for Share Secret Key Generation // IEEE Transactions on Information Forensics and Security. 2016. Vol. 11. Iss. 12. PP. 2691–2705. DOI:10.1109/TIFS.2016.2594143

31. Yakovlev V., Korzhik V., Starostin V., Lapshin A. Channel Traffic Minimizing Key Sharing Protocol Intended for the Use over the Internet and Secure without any Cryptographic Assumptions // Proceedings of the 32nd Conference of Open Innovations Association FRUCT (Tampere, Finland, 9–11 November 2022). FRUCT 32, 2022. PP. 300–307. DOI:10.23919/FRUCT56874.2022.9953895

32. Lai E., Gamal H.El., Poor H.V. The Wiretap Channel with Feedback Encryption over the Cannel // IEEE Transactions on Information Theory. 2008. Vol. 54. Iss. 11. PP. 5059–5067. DOI:10.1109/TIT.2008.929914

33. Korzhik V., Yakovlev V., Starostin V., Lapshin A. Vulnerability of the Key Sharing Protocol Executing over the Noiseless Public Cannels with Feedback // Proceedings of the 35th Conference of Open Innovations Association FRUCT-35 (Tampere, Finland, 24–26 April 2024). FRUCT-35, 2024. PP. 374–379 DOI:10.23919/FRUCT61870.2024.10516344

References

1. Alpern B., Schneider F.B. Key exchange using 'keyless cryptography'. *Information Processing Letters*. 1983:16(2):79–81. DOI:10.1016/0020-0190(83)90029-7

2. Korzhik V. Keyless cryptography. *Proceedings of the 9th International Conference on System Administration, Networking and Security, Orlando, USA*. 2000

3. Korzhik V., Bakin M. Information-theoretical Secure Keyless intensification. *Proceedings of the 2000 IEEE International Symposium on Information Theory, 25–30 June 2000, Sorrento, Italy*. Piscataway: IEEE Press; 2000. DOI:10.1109/ISIT.2000.866310

4. Korzhik V. Keyless cryptography. *Invited Talk at Security Seminar at CERIAS Purdue University*. 2001

5. Mukherjee A., Fakoorian S.A.A., Huang J., Swindlehurst A.L. Principles of Physical Layer Security in Multiuser Wireless Network A. Survey. *IEEE Communications Surveys & Tutorials*. 2014;16(3):1550–1573. DOI:10.1109/SURV.2014.012314.00178

6. Wyner A.D. The Wire-tap channel. *Bell System Technical Journal*. 1975;54(8):1355–1387. DOI:10.1002/j.1538-7305.1975.tb02040.x

7. Bennett C.H., Bessette F., Brassard G., Salvail L., Smolin J. Experimental quantum cryptography. *Journal of Cryptology*. 1992;5:3–28. DOI:10.1007/BF00191318

8. Kushnir D.V. *Research and Development of Methods of Confidential Data Distribution on Quanto Channels*. PhD Thesis. St. Petersburg: The Bonch-Bruevich Saint-Petersburg State University of Telecommunications Publ.; 1996. (in Russ.)

9. Korzhik V.I., Yakovlev V.A. Fundamentals of Cryptology St. Petersburg: Intermediia Publ.; 2016. 296 p. (in Russ.) EDN:WEQWMN

10. Korzhik V., Kushnir D. Key sharing based on the wire-tap channel type ii concept with noisy main channel // Proceedings of the International Conference on the Theory and Applications of Crypotology and Information Security, ASIACRYPT '96, 3–7 November 1996, Kyongju, Korea. Lecture Notes in Computer Science, vol.1163. Berlin, Heidelberg: Springer; 1996. p.210–217. DOI:10.1007/BFb0034848

11. Liu Y., Zhang W.J., Jiang C., Chen J.P., Zhang C., Pan W.X., et al. Experimental Twin-Field Quantum Key Distribution over 1000 km Fiber Distance. *Physical Review Letters*. 2023;130:210801. DOI:10.1103/PhysRevLett.130.210801

12. Milov M., Pham T.M., Chorti A., Barreto A.N., Fettweis G. Physical Layer Security – From Theory to Practice. *IEEE BITS the Information Theory Magazine*. 2023;3(2):67–79. DOI:10.1109/MBITS.2023.3338569

13. Shannon K.E. *Works on Information Theory and Cybernetics*. Moscow: Foreign Literature Publ.; 1963. 829 p. (in Russ.)

14. Maurer U.M. Secret key agreement by public discussion based on common information. *IEEE Transactions on Information Theory*. 1993;39(3):733–742. DOI:10.1109/18.256484

15. Korzhik V.I., Yakovlev V.A. Non-asymptotic estimation of the efficiency of coded noise cancellation in channels with diversion. *Problems of Information Transmission.* 1991;4:223–228. (in Russ.)

16. Peterson W., Welnod E. *Codes, Fixed Errors*. Moscow: Mir Publ.; 1976. (in Russ.)

17. Korzhik V.I., Yakovlev V.A. Information protection from leakage due to side electromagnetic radiation and interference based on the code noise method. *Informatika i vychislitelnaia tekhnika*. 1993;8(1-2):61–66. (in Russ.)

18. Korzhik V., Morales-Luna G., Balakirsky V.B. Privacy amplification theorem for noisy main channel. *Proceedings of the 4th International Conference on Information Security, ISC 2001, 1–3 October 2001, Malaga, Spain. Lecture Notes in Computer Science, vol.2200*. Berlin, Heidelberg: Springer; 2001. p.18–26. DOI:10.1007/3-540-45439-X_2

19. Yakovlev V., Korzhik V. Key Distribution Protocol Based on Noisy Channels in Presence of Active Adversary. *IEEE Transactions on Information Theory*. 2008;54(6):2535–2550. DOI:10.1109/TIT.2008.921689

20. Schneier B. *Applied Cryptography*. Moscow: Triumf Publ.; 2002. (in Russ.)

21. Korzhik V., Starostin V., Yakovlev V., Kabardov M., Krasov A., Adadurov S. Advance in Keyless Cryptography. Chapter 6. *In: Ramakrishnan S. (ed.) Lightweight Cryptographic Techniques and Cybersecurity Approaches*. 2022. p.97–117. DOI:10.5772/intechopen.104429

22. Tilborg H.C.A. *Encyclopedia of Cryptography and Security*. Springer, 2005.

23. Myasnikov A.G., Shpilrain V., Ushakov A. *Non-commutative Cryptography and Group-theoretic Problems*. American Mathematical Society; 2011. 385 p.

24. Goldreich O., Goldwasser S., Halevi S. Public-key cryptosystems from Lattice reduction problems. *Proceedings of the 17th Annual International Cryptology Conference, CRYPTO '97, 17–21 August 1997, Santa Barbara, USA. Lecture Notes in Computer Science, vol.1294*. Berlin, Heidelberg: Springer; 1997. p.112–131. DOI:10.1007/BFb0052231

25. Dean T., Goldsmith Aj. Physical layer cryptography through massive MIMO. *Proceedings of the 2013 IEEE Information Theory Workshop, ITW, Seville, Spain, 9–13 September 2013*. Piscataway: IEEE; 2013. p.1–3. DOI:10.1109/ITW.2013.6691222

26. Ben-Israel A., Greville T.N.E. *Generalized Inverses: Theory and Applications*. Springer; 2003.

27. Steinfeld R., Sakzad A. On massive MIMO physical layer cryptosystems. *Proceedings of IEEE Information Theory Workshop – Fall, ITW, 11–15 October 2015, Jeju, Korea (South)*. IEEE; 2015. p.292–296. DOI:10.1109/ITWF.2015.7360782

28. Aono T., Higuchi K., Ohira T., Komiyama B., Sasaoka H. Wireless secret key generation exploiting reactance-domain scalar response of multipath fadiny cannels. *IEEE Transactions on Antennas and Propagation*. 2005;53(11):3776–3784. DOI:10.1109/TAP.2005.858853

29. Korzhik V., Yakovlev V., Kovajkin Y. Secret Key Agreement Over Multipath Channels Exploiting a Variable-Directional Antenna. *International Journal of Advanced Computer Science and Applications*. 2012;3(1):172–178.

30. Qin D., Ding Z. Exploiting Multi Antenna Non-Reciprocal Channels for Share Secret Key Generation. *IEEE Transactions on Information Forensics and Security*. 2016;11(12):2691–2705. DOI:10.1109/TIFS.2016.2594143

31. Yakovlev V., Korzhik V., Starostin V., Lapshin A. Channel Traffic Minimizing Key Sharing Protocol Intended for the Use over the Internet and Secure without any Cryptographic Assumptions. *Proceedings of the 32nd Conference of Open Innovations Association FRUCT, 9–11 November 2022, Tampere, Finland*. FRUCT-32; 2022. p.300–307. DOI:10.23919/FRUCT56874.2022.9953895

32. Lai E., Gamal H.El., Poor H.V. The Wiretap Channel with Feedback Encryption over the Cannel. *IEEE Transactions on Information Theory*. 2008;54(11):5059–5067. DOI:10.1109/TIT.2008.929914

33. Korzhik V., Yakovlev V., Starostin V., Lapshin A. Vulnerability of the Key Sharing Protocol Executing over the Noiseless Public Cannels with Feedback. *Proceedings of the 35th Conference of Open Innovations Association FRUCT-35, 24–26 April 2024, Tampere, Finland*. FRUCT Oy; 2024. p.374–379 DOI:10.23919/FRUCT61870.2024.10516344

Статья поступила в редакцию 29.10.2024; одобрена после рецензирования 21.11.2024; принята к публикации 09.12.2024.
The article was submitted 29.10.2024; approved after reviewing 21.11.2024; accepted for publication 09.12.2024.

Информация об авторах:

КОРЖИК Валерий Иванович
доктор технических наук, профессор, профессор кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0002-8347-6527

ЯКОВЛЕВ Виктор Алексеевич
доктор технических наук, профессор, профессор кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0009-0007-2861-9605

СТАРОСТИН Владимир Сергеевич
кандидат физико-математических наук, доцент, доцент кафедры высшей математики Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0009-0000-2939-1971

БУЙНЕВИЧ Михаил Викторович
доктор технических наук, профессор, профессор кафедры прикладной математики и информационных технологий Санкт-Петербургского университета ГПС МЧС России
https://orcid.org/0000-0001-8146-0022

Коржик В.И. и Буйневич М.В. являются членами редакционного совета журнала «Труды учебных заведений связи» с 2016 г., но не имеют никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.

Korzhik V.I. and Buinevich M.V. have been a members of the journal "Proceedings of Telecommunication Universities" Editorial Council since 2016, but have nothing to do with the decision to publish this article. The article has passed the review procedure accepted in the journal. The authors have not declared any other conflicts of interest.

Научная статья
УДК 004.5
https://doi.org/10.31854/1813-324X-2024-10-6-99-110

Система статистического измерения атомарной эффективности графических элементов интерфейсов

Курта Павел Андреевич, expert@kurta.ru

Санкт-Петербургский университет ГПС МЧС России,
Санкт-Петербург, 196105, Российская Федерация

Аннотация

Актуальность. *В настоящее время качество интерфейсов зачастую имеет решающую роль в решении задач человеком с применением информационных сервисов. Для оценки интерфейсов ранее было введено понятие эффективности, состоящей из следующих показателей: результативности – условной меры количества ошибок при работе с информационной системой; оперативности – скорости работы пользователя с информационной системой для получения требуемого результата; ресурсоэкономности – степени психоэмоционального напряжения пользователя при вводе и обработке данных. Тем не менее, полученная ранее модель требует не только применения математического аппарата для определения таких показателей, но и знаний об атомарных (т. е. частных или обособленных) эффективностях графических элементов, связанных с особенностями взаимодействия с ними пользователя.*

Целью *настоящей статьи является повышение эффективности интерфейсов информационных сервисов, для чего требуется вычисление атомарных эффективностей отдельных графических элементов.*

Сущность *предлагаемого решения заключается в визуальной системе статистического измерения атомарных эффективностей шести графических элементов (текстового поля, выпадающего списка, классической и флаговой кнопки, двунаправленного счетчика и «ползунка») по результатам выполнения с их помощью различных заданий пользователями. Так, например, оперативность текстового поля по сравнению с выпадающим списком будет выше для коротких слов – поскольку их ввод с клавиатуры быстрее выбора из списка, но ниже для длинных предложений – в этом случае человеку быстрее выбрать нужное, чем обеспечивать корректный ввод. Принцип измерения эффективностей предложенной системой* ***основан*** *на последовательном выводе графических форм с различным типом (в ряде случае – и количеством) элементов, указании заданий пользователю и измерению корректности и времени ввода данных. Для снижения субъективности в действиях применяются различные приемы, такие, как таймеры различной длительности. Для оценки психоэмоционального напряжения используется специальный опрос в конце групп тестов элементов. Система* ***имеет реализацию*** *в виде Web-сайта на языке PHP, отдельные Web-страницы которого и их интерпретация приведены в статье.* ***Эксперименты*** *с применением данной системы для 50 пользователей позволили получить искомые атомарные эффективности всех элементов.*

Научная новизна *решения состоит в возможности получения оценок эффективности элементов интерфейса полностью формальным способом, учитывающим только особенности взаимодействия с ними пользователей, а также специфику данных (размер, тип).*

Теоретическая значимость *состоит в расширении класса способов оценки эффективности графических интерфейсов, полученной через измерение характеристик составляющих его элементов.*

Практическая значимость *заключается в возможности непосредственного применения полученных графиков атомарных эффективностей для сравнения интерфейсов и их оптимизации.*

Ключевые слова: *интерфейс, графический элемент, эффективность, статистика, измерение, прототип, эксперимент*

Ссылка для цитирования: Курта П.А. Система статистического измерения атомарной эффективности графических элементов интерфейсов // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 99–110. DOI:10.31854/1813-324X-2024-10-6-99-110. EDN:VONRKD

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-99-110

System for Statistical Measurement of Atomic Efficiency for Graphical Interface Elements

Kurta Pavel A., expert@kurta.ru

Saint Petersburg University of State Fire Service of Emercom of Russia,
St. Petersburg, 190000, Russian Federation

Annotation

Relevance*. Currently, the interfaces quality often plays a decisive role in solving problems by a person using information services. To evaluate interfaces, the efficiency concept was previously introduced, consisting of the following indicators: effectiveness – a conditional errors number measure when working with an information system; efficiency – the speed of the user's work with the information system to obtain the desired result; 3) resource efficiency – the degree of user psycho-emotional stress when entering and processing data. Nevertheless, the previously obtained model requires not only the mathematical apparatus usage to determine such indicators, but also knowledge of the graphic elements atomic (i.e. individual or isolated) efficiencies associated with the features of the user's interaction with them.*

The *article* ***purpose*** *is to improve the efficiency of information service interfaces, which requires calculating the individual graphic elements atomic efficiencies.*

The *proposed solution* ***essence*** *is the visual system for statistical measurement of the atomic efficiency for 6 graphic elements (text field, drop-down list, classic and checkbox button, bidirectional counter and "slider") based on the performing various tasks results by users with their help. For example, the text field efficiency compared to a drop-down list will be higher for short words – since their input from the keyboard is faster than selecting from the list, but lower for long sentences – since in this case it is faster for a person to select the desired one than to provide correct input. The measuring principle the efficiency by the proposed system is* ***based on*** *the sequential output of graphic forms with different types (and in some cases, quantities) of elements, indicating the task to the user and measuring the correctness and data entry duration. To reduce subjectivity in actions, various techniques are used, such as different duration timers. A special survey at the end of the element test groups is used to assess the psycho-emotional load. The system* ***has an implementation*** *as a Web site in PHP, individual Web pages of which and their interpretation are given in the article.* ***Experiments*** *with the use of this system for 50 users allowed us to obtain the desired all elements atomic efficiencies.*

The scientific novelty *of the solution lies in the obtaining possibility estimates of the interface elements efficiency in a completely formal way, taking into account only the features of user interaction with them, as well as the data specifics (size, type).*

The theoretical significance *lies in expanding the class of methods for assessing the graphical interfaces efficiency obtained through assessing the elements that make it up.*

The practical significance *lies in the possibility of directly using the obtained atomic efficiencies graphs to compare interfaces and optimize them.*

Keywords: *interface, graphic element, efficiency, statistics, measurement, prototype, experiment*

For citation: Kurta P.A. System for Statistical Measurement of Atomic Efficiency for Graphical Interface Elements. *Proceedings of Telecommunication Universities.* 2024;10(6):99–110. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-99-110. EDN:VONRKD

Введение

Информационные технологии (далее – ИТ) стали неотъемлемой частью современного мира, одним из предназначений которых является удовлетворение информационных потребностей общества. Так, например, для заказа и оплаты железнодорожных билетов уже нет необходимости обращаться в кассы, а созданы и успешно эксплуатируются соответствующие программные терминалы; или же для поиска местоположения магазинов в торговых центрах есть необходимые электронные справочные. Таким образом, человек в современном обществе

окружен огромным количеством информационных потоков, взаимодействие с которыми (восприятие, обработка и генерация новых) уже стало необходимым условием существования. Это, в свою очередь, требует высокой эффективности от основных элементов, обеспечивающих такое взаимодействие, а именно – от интерфейсов (в данном случае, в первую очередь естественно понимаются текстовые и графические, а не программные) [1]. С другой стороны, разнородность решаемых информационных задач (например, управление беспилотными летательными аппаратами [2]), сложность проектирования соответствующих систем, а также недостаточная развитость и применение стандартов области приводят к тому, что интерфейсы взаимодействия человека с информационной системой (далее – ИС) создаются организациями с использованием собственных представлений (зачастую, сложившихся исторически или эмпирически) об их эффективностях [3]. Таким образом, требуется научно-обоснованное определение эффективности интерфейсов взаимодействия, а также создание путей ее повышения – что является основной целью исследования автора (которое частично будет описано далее). В данной работе описывается один из этапов такого исследования, а именно определение измерения эффективности отдельных графических элементов интерфейса, названное *атомарной эффективностью*.

Результаты исследования

Поскольку текущий этап (притом непоследний) является частью более крупного исследования, то кратко укажем как полученные ранее результаты, так и планируемые в будущем.

Предпосылки

Исследование современного состояния дел в области проектирования интерфейсов и их применения пользователями позволили построить общую схему такого взаимодействия, а также ряд возможных дефектов в таких системах (например, «Тупиковый путь сценария» или «Игнорирование ограничений организма человека») [4]. Был выделен отдельный класс широко востребованных информационных сервисов, названных запросным и предназначенных для решения задач, инициируемых пользователем и основанных на вводимых им данных (например, поисковая ИС) [5].

Эффективность интерфейса

В результате дальнейшего анализа была предложена и введена эффективность интерфейса ($Efficiency$), состоящая из трех следующих компонент: 1) результативность ($Potency$) – условная мера количества ошибок, которые может допустить пользователь при работе с ИС; 2) оперативность ($Operativeness$) – скорость (как обратной времени) работы пользователя с ИС для получения требуемого результата; 3) ресурсоэкономность (*Resource-Saving*) – степень психоэмоционального напряжения (далее – ПЭН) пользователя, которое было вызвано вследствие его нагрузки при вводе данных и обработке результатов [6].

Модель

Затем, используя предварительно полученные результаты, формализация взаимодействия с интерфейсом позволила описать единую эффективностную модель процесса, суть которой состояла в следующем [7]. Логика решения задач пользователя с помощью интерфейса является линейной – от пошагового ввода данных до вывода конечного результата. Весь интерфейс представляется как последовательность форм с размещенными на них графическими элементами для ввода и вывода данных; также выделены 4 следующих базовых класса элементов [8]: текстовое поле, выпадающий список, флаговая кнопка (т. е. «чекбокс») и двунаправленный счетчик (т. е. «спиннер»). Переходы между графическими формами осуществляются путем нажатия на соответствующий элемент-кнопку; последняя форма является завершающей логику решения задачи и не содержит кнопки. При этом все графические элементы и формы имеют собственную эффективность – атомарную, вносящую основной вклад в общую эффективность всего интерфейса. При этом для элементов атомарная эффективность определяется как их способностью по передаче данных человеку (в обе стороны), так и самими данными (по крайне мере, размером); так, например, использование выпадающего списка для ввода большого целого значения, очевидно, будет менее «удачным» решением с точки зрения оперативности работы, чем использование текстового поля (хотя в случае последнего может возрасти количество ошибок ввода – т. е. уменьшится результативность). Для графических форм их атомарная эффективность определяется количеством элементов на них, поскольку «перенасыщение» источниками и получателями информации может, некоторым образом, сказываться на работе с формой человека. Также атомарные эффективности элементов и форм должны корректироваться с учетом их дальности расположения в логике решения задачи таким информационным сервисом, т. к. работа пользователя с поздними элементами будет чуть более затрудненной, чем с первичными (хотя бы из-за небольшого роста ПЭН). И хотя описанная модель позволяет теоретически оценивать эффективность всего интерфейса полностью формальным способом, тем не менее для практического применения требуется получение конкретных значений атомарных эффективностей всех основных классов элементов и форм, а также определение влияния на нее их положения в логике сервиса. При этом уже сейчас можно предположить, что все ато-

марные эффективности будут представлять не конкретные числа, а некоторые аналитические зависимости от особенностей вводимых и выводимых данных; по крайне мере, от их размера. Также основное влияние на общую эффективность интерфейса будут иметь именно эффективности элементов, как основных точек взаимодействия человека с данными ИС.

Система статистического измерения

Для определения атомарных эффективностей графических элементов была построена соответствующая система статистического измерения (далее – Система), описанию и экспериментам с которой и посвящен текущий исследовательский этап (и, соответственно, статья).

Метод оптимизации

Получение всех атомарных эффективностей (элементов, форм), а также формул для их корректировки с учетом положения в логике решения задачи позволит оценивать любой интерфейс взаимодействия пользователя с информационным сервисом запросного типа, что, хотя и можно считать важным научно-практическим результатом, но оно, однако, не позволит решить основную цель исследования – улучшить сами интерфейсы путем повышения их интегральной эффективности (как уже для созданных, так и еще проектируемых). Для этого потребуется оптимизация архитектуры интерфейса – как совокупности форм и расположения элементов на них. Для этого требуется создание соответствующих методов оптимизации, используемых в качестве целевой функции эффективностной модели (т. е. результаты ее применения), максимизацию которой необходимо обеспечить. Исходя из сути задачи, подходящими методами можно считать комбинаторные, для реализации которых (не без оснований следуя современных трендам развития ИТ-области) может быть применен искусственный интеллект в части генетических или подобных им алгоритмов [9].

Обзор работ

Проведем далее общий обзор работ, посвященных существующим решениям (теоретическим и практическим) по оценке характеристик графических элементов.

Фундаментальная монография [10] посвящена оценке визуальной составляющей графических элементов, к которой, в частности, относятся используемые цвета, композиция, информативность и персонализация. Также одним из авторов монографии в [11] качественно оценивается применимость различных методов и метрик к различным показателям эффективности эргономической эстетики интерфейса.

В [12] предлагается методика оценки качества интерфейсов с позиции оперативности (относительно элементарных и целевых операций) на примере пунктов централизованной охраны. В частности, интерфейс автоматизированного рабочего места делится на множество экранных форм, предназначенных для решения определенной задачи; каждая же форма состоит из набора графических элементов.

В статье [13] оценку удобства интерфейсов предлагается производить одним из следующих методов: опросный, экспертный, аналитический и экспериментальный; для каждого из которых указаны преимущества и недостатки.

Исследователи в [14] предлагают оценивать психоэмоциональное состояние (очевидно, коррелирующее с ПЭН) пользователя путем непосредственного внедрения соответствующего модуля в сам интерфейс. Источником состояния в данном случае будет изображение лица пользователя с камеры (например, на телефоне), эмоция которого затем будет программно распознаваться.

В работе [15] для оценки интерфейса предлагается использовать искусственную нейронную сеть (далее – ИНС) и метод анализа иерархий (далее – МАИ), основанные на субъективных ожиданиях пользователя. Для этого, в том числе, требуется подбор эталонных субъектов проверки (т. е. «валидаторов» данных) и выделение индексов оценки (например, согласование цветов, удобство ввода, частота системных ошибок и время отклика). Показано, что результаты работы ИНС согласуются с аналогичными результатами, полученными с помощью МАИ.

Как показал краткий, но показательный обзор работ, приводимые в них решения обладают, как правило, высокой степенью субъективности при оценке характеристик интерфейсов. При этом, практически отсутствуют модели, обладающие достаточной адекватностью (т. е. соответствием реальному объекту моделирования – интерфейсу). С этой позиции, хотя описанная ранее авторская модель также и содержит субъективные элементы (атомарную эффективность и ее корректировки), тем не менее, последние могут быть получены лишь один раз, а затем использоваться как справочный материал, отражающий особенности взаимодействия человека с теми или иными частями графического интерфейса.

Описание системы

Дадим описание основной идеи и реализации разработанной Системы.

Идея

Основная идея Системы заключается в измерении заданных характеристик действий пользователей с интерфейсными элементами. По результатам измерений будут вычислены показатели атомарной эффективности элементов: 1) результативности – путем подсчета количества сделанных при вводе

ошибок; 2) оперативности – обратной времени, затраченного пользователем; 3) ресурсоэкономности – субъективной оценке пользователя касательно сложности выполнения действий.

Архитектура

Программная реализация Системы представляет собой Web-сайт с последовательностью страниц, на каждой из которой пользователю описывается задание для ввода данных (числовых и текстовых), а также предоставляется для этого один из нескольких следующих классов элементов интерфейса: четырех упомянутых ранее и двух дополнительных – классической кнопки и «ползунка». Выполнение действий осуществляется в различных условиях и ограничениях, задаваемых количеством элементов, объемом выводимых данных, допустимым временем работы и т. п. По мере выполнения заданий на каждой странице Система измеряет все необходимые характеристики пройденных тестов, записывая результат в базу данных. Также вначале работы пользователь указывает данные о себе, позволяющие вычислять атомарные эффективности для разных возрастных групп и сфер деятельности.

Группы тестов

Основные группы тестов (определяемые Web-формами на визуальных страницах Системы и, соответственно, сохраняемыми в базе данных результатами) состоят из тестирования:

– ввода текста и чисел в текстовое поле;

– выбора текста и чисел в выпадающем списке;

– ввода текста и чисел с помощью флаговой кнопки;

– ввода текста и чисел с помощью классической кнопки;

– выбора числа с помощью двунаправленного счетчика;

– выбора числа с помощью «ползунка».

Таким образом, общее число групп тестов для различных графических элементов и типов данных составило 10 (а не $6 \times 2 = 12$, поскольку двунаправленный счетчик и ползунок позволяют вводить только числа).

В первой паре групп тестов длина данных увеличивалась от 1 символа до 10. Во второй паре групп тестов менялось количество кнопок на форме, затрудняя тем самым ввод пользователю. В третьей, последней, паре групп тестов оценивался ввод только чисел, задаваемых диапазоном от 1 до максимального значения, поддерживаемого элементом. Оценка ПЭН осуществлялась субъективной пользователем через соответствующую форму Системы.

Примеры Web-страниц

Примеры отображаемых форм для ввода числовых и текстовых значений, наглядно демонстрирующие принцип функционирования разработанной Системы, представлены далее.

На рисунке 1 пользователю предлагается ввод в текстовое поле английской комбинации букв «idgq»; соответственно, после ввода значения и нажатия кнопки «Далее» будет посчитано одно из статистических значений атомарной эффективности элемента текстового поля – исходя из того, верное ли значение ввел пользователь, и сколько он затратил на это время. В такой схеме расчета присутствует очевидная проблема, заключающаяся в том, что любой человек, скорее всего, потратит все доступное время на обеспечение корректности ввода; как следствие, результативность элемента будет высчитываться всегда максимальной – в ущерб оперативности. Поэтому для недопущения такого эффекта «чрезмерной ответственности» форма имеет таймер времени, уменьшение которого до 0 секунд будет означать ошибку ввода (т. е. резкое снижение результативности). Такое ограничение, в свою очередь, «подстегнет» пользователя укладываться в отведенное для ввода время и позволит более адекватно оценивать эти показатели эффективности. У данной формы есть несколько вариаций, заключающихся в постепенном увеличении длины текста для ввода и уменьшении отводимого для этого времени, что позволит собрать больший объем статистических данных касательно эффективности текстового поля в различных условиях функционирования. Также присутствует кнопка «Пауза» для приостановки тестирования; при этом текущая форма пропадает, не давая тем самым пользователю выбрать нужный вариант текста при остановленном счетчике времени выполнения задания – т. е. не допуская «обмана» Системы. Уменьшающийся таймер времени, как и кнопка приостановки, присутствуют на каждой из форм ввода и далее указываться не будут.

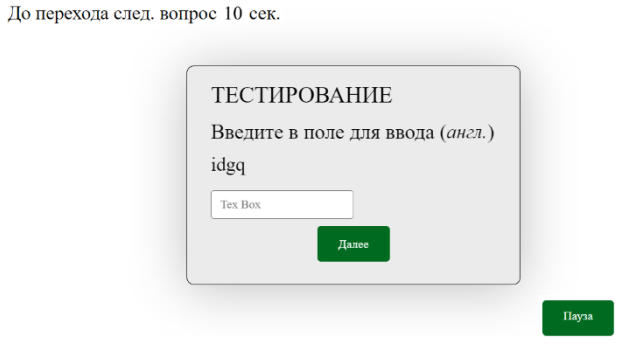


Рис. 1. Пример формы для оценки атомарной эффективности текстового поля (с таймеров времени и кнопкой приостановки)

Fig. 1. Example of a Form for Evaluating the Text Field Atomic Efficiency (with Timer and Pause Button)

Данный элемент (т. е. текстовое поле), по сути, можно считать первым появившимся в интерфейсах, поскольку под ним условно понимается любая командная строка (или консоль) большинства Unix-подобных операционных систем.

На рисунке 2 от пользователя требуется выбор определенного текстового значения из выпадающего списка.

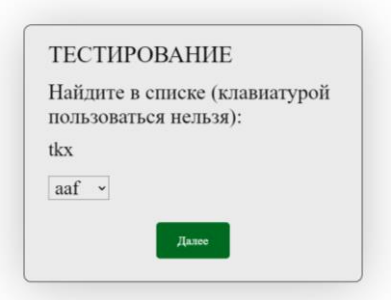


Рис. 2. Пример части формы для оценки атомарной эффективности выпадающего списка

Fig. 2. Example of the Form Part for Evaluating the Drop-Down List Atomic Efficiency

Уже сейчас можно предположить, что использование данного интерфейсного элемента будет более предпочтительным для небольшого количества текстовых строк большой длины; в ином случае (много текстовых строк малой длины, например, выбор инициалов) большей результативностью и оперативностью будет обладать классический ввод в текстовом поле.

Немного иной подход был применен для вычисления атомарной эффективности элемента-кнопки (рисунок 3). Для этого на форме в случайном порядке размещалось определенное количество кнопок с текстовыми метками, одна из которых являлась целевой – нажатие на нее считалось верным выполнением текущего задания.

Рис. 3. Пример части формы для оценки атомарной эффективности кнопок

Fig. 3. Example of the Form Part for Evaluating the Buttons Atomic Efficiency

Естественно, в случае одной кнопки, ее результативность и оперативность вычислялись, как максимальная (поскольку от пользователя требовалось лишь нажать на единственный элемент на форме). Однако при увеличении количества кнопок пользователь начинал затрачивать время на поиск нужной кнопки, имея при этом некоторый шанс ошибиться – атомарная эффективность начинала приобретать нетривиальный характер.

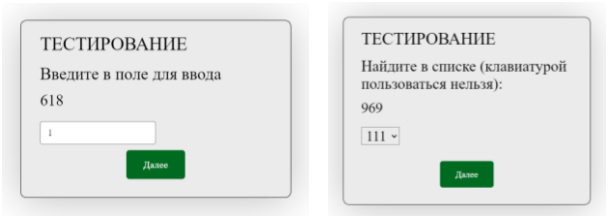
Аналогичным образом осуществляется статистическая оценка атомарной эффективности флаговой кнопки, что показано на рисунке 4.

Рис. 4. Пример части формы для оценки атомарной эффективности флаговой кнопки

Fig. 4. Example of the Form Part for Evaluating the Checkbox Atomic Efficiency

Пользователю также необходимо (естественно, как и ранее, в отведенное время) выбрать элемент с заданной текстовой меткой.

Тестирование ввода числовых данных осуществляется полностью аналогичным образом, что представлено на рисунке 5.



a) b)

Рис. 5. Пример части формы для оценки атомарной эффективности a) текстового поля и b) выпадающего списка

Fig. 5. Example of the Form Part for Evaluating the Atomic Efficiency of a Text Field (a) and a Drop-Down List (b)

Длина данных, которые требовалось ввести на различных формах, составляла от 1 до 10, что охватывает достаточно большой спектр возможных задач, решаемых с помощью информационных сервисом запросного типа (например, поиск человека по фамилии и дате рождения).

Также в начале тестирования у пользователя запрашивался его возраст с помощью следующих градаций:

– детство (менее 10 лет);
– отрочество (10–16 лет);
– юность (17 лет –24 года);

– молодость (25 лет – 44 года);
– зрелость (45–60 лет);
– старость (более 60 лет).

Естественно, хотя возрастные группы и были выбраны условно, тем не менее, они отражают основные этапы становления человека, качественно влияющие и на его способности к работе с интерфейсами.

За формой с возрастом пользователя следовал ввод сферы деятельности из следующего списка: «Бизнес», «Научная», «Творческая», «Военная», «Государственная», «Другая» (в том числе обучающиеся). Такие данные о пользователе позволяют не просто определить атомарные эффективности всех элементов, но и персонифицировать для соответствующих возрастных и деятельностных групп. Данное разделение эффективностей обосновывается тем, что по мере взросления и, исходя из специфики повседневной деятельности, адаптированность пользователей к взаимодействию с интерфейсами гипотетически должна меняться (что многократно можно наблюдать в повседневной жизни). Например, молодое поколение более оперативно взаимодействует с формами, перегруженными информационными элементами, чем зрелое, а сотрудники творческой сферы могут допускать ошибки с теми элементами, где ценящие точность люди бизнеса будут максимально аккуратны.

В конце этапов тестирования на различных элементах проводился опрос пользователя касательно его ПЭН в виде указания степени усталости по 5-балльной системе (от «не устал» до «очень устал»).

Представление результатов

Пример результирующих характеристик пройденных тестов одного пользователя, сохраненных в базе данных Системы, представлен на рисунке 6 (имя, фамилия, возраст и сфера деятельности были преднамеренно скрыты).

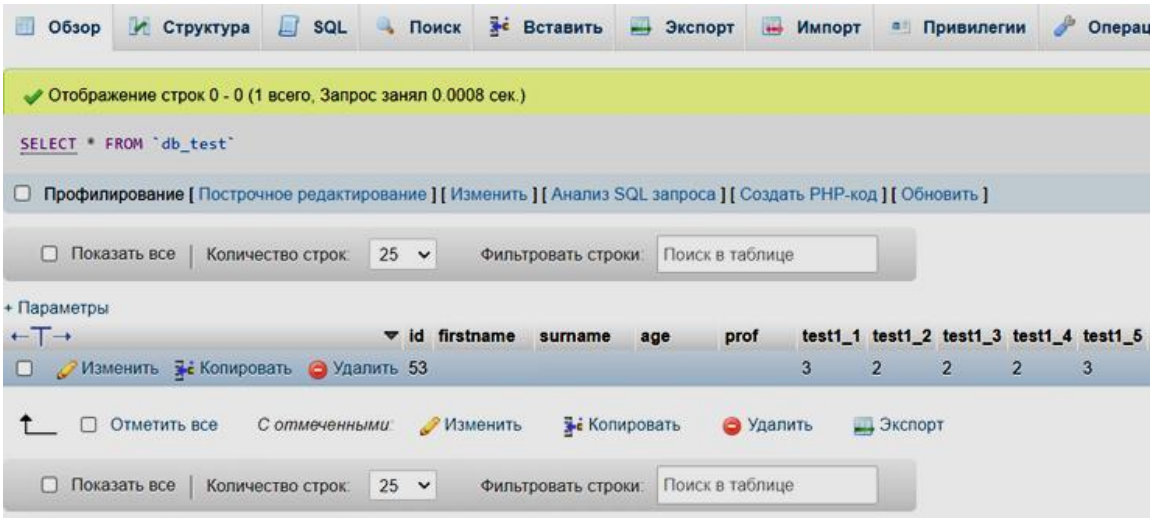


Рис. 6. Пример характеристик пройденных тестов пользователем в базе данных

Fig. 6. Example of the Tests Characteristics Passed by a User in the Database

Результат прохождения каждой формы записывался в виде числового значения (см. рисунок 6) со следующей интерпретацией:
0 – была допущена ошибка при вводе;
–1 – окончилось время ввода (т. е. таймер дошел до 0 ранее, чем пользователь нажал кнопку «Далее»);
положительное число – время (в секундах), затраченное пользователем на ввод данных.

Эксперимент

Входные данные

В качестве фокус-группы, на которой проводилось тестирование Системой, было отобрано 50 студентов из Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ), обучающихся на бакалавриате с разных факультетов, не знакомых друг с другом и имеющих различные средние оценки по успеваемости. Таким образом, была обеспечена достаточно разнородная выборка пользователей оцениваемых интерфейсных элементов Системы, относящаяся к возрастной группе «Юность» и сферы деятельности «Другая».

Таблица с результатами

Часть результирующих числовых значений, измеренных Системой и сведенных в Excel-таблицу, приведены на рисунке 7 в виде тепловой карты; используются следующие цвета:

– темно-красный (для значения «–1»; т. е. нехватки времени);

– светло-красный (для значения «0»; т. е. ошибки при вводе данных);

– зеленый (для значений больше или равных 1; т. е. время выполнения задания в секундах).

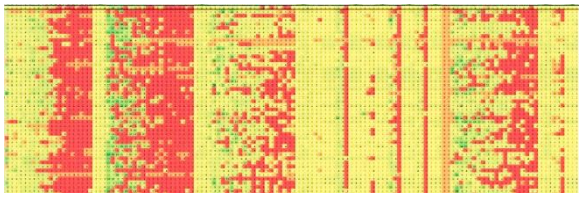


Рис. 7. Тепловая карта результатов измерения атомарных эффективностей графических элементов

Fig. 7. Heat Map of the Measuring Results for Graphic Elements Atomic Efficiencies

Тепловая карта результатов (см. рисунок 7) достаточно хорошо отражает закономерности при выполнении заданий всеми участниками эксперимента – периодически появляющиеся красные вертикальные волны соответствуют вариантам тестов с малым значением таймера времени на его выполнение, что приводит к практически невозможному завершению задания пользователями. Также, как можно видеть, аномалии в результатах наблюдаются крайне редко; например, полностью светло- или темно-красными линии, означающие «нечестность» прохождения заданий, для случаев, когда пользователь или долгое время не осуществлял действий (заканчивалось время) или выбирал первое попавшееся решение (результат был всегда неверным). Таким образом, по крайне мере, визуально, статистически собранную информацию можно считать корректной.

Обработка результатов

В результате обработки части результатов (для этого, в частности, выбирался ввод данных длиной от 1 до 10 символов) была получена возможность построения зависимости показателей атомарной эффективности каждого элемента от параметров обрабатываемых данных с использованием статистических данных касательно работы с ним пользователя, представленная затем в виде соответствующего графика.

Приведем формулы расчета всех показателей эффективности с учетом текущей реализации Системы. Также отметим, что под размером данных (которые было необходимо ввести пользователю) в Системе понимаются различные значения, зависящие от тестируемого элемента: для текстового поля – длина строки, для выпадающего списка – количество его элементов, для флаговой кнопки и классической кнопки – их количество на экране, для двунаправленного счетчика и «ползунка» – количество задаваемых с помощью их значений.

Далее в аналитических записях тип данных и их размер будем указывать с помощью двух верхних индексов:

$$\begin{cases} Type \in \{Text, Integer\} \\ Length \in [1 \dots 10] \end{cases},$$

где $Text$ и $Integer$ – указание типа данных, как текстовый и числовой, соответственно.

Значения всех показателей для тестов с различными значениями таймеров усреднялись и для упрощения записи дальше будут указываться в итоговом виде.

Показатель оперативности вычислялся по следующей формуле:

$$\begin{cases} Operativeness^{Type,Length} = \dfrac{1}{Time^{Type,Length}} \\ Time^{Type,Length} = \dfrac{\sum_i Time_i^{Type,Length}}{N} \end{cases},$$

где $Time^{Type,Length}$ – среднее время (в секундах) выполнения действий пользователя по вводу данных типа $Type$ и длины $Length$; $Time_i^{Type,Length}$ – аналогичное время ввода i-м пользователем; N – количество всех пользователей.

Таким образом, показатель оперативности представляет собой усредненную скорость работы с элементом для всех пользователей.

Показатель ресурсоэкономности вычислялся по следующей формуле:

$$\begin{cases} ResourceSaving^{Type,Length} = \dfrac{5 - Feeling^{Type,Length}}{4} \\ Feeling^{Type,Length} = \dfrac{\sum_i Feeling_i^{Type,Length}}{N} \\ Feeling_i^{Type,Length} \in [1 \dots 5] \end{cases},$$

где $Feeling^{Type,Length}$ – средняя субъективная оценка ПЭН (в баллах от 1 для «не устал» до 5 для «очень устал»), данная пользователем после выполнения действий по вводу данных типа $Type$ и длины $Length$; $Feeling_i^{Type,Length}$ – аналогичная оценка ПЭН, данная i-м пользователем.

Таким образом, показатель ресурсоэкономности представляет собой усредненную оценку ПЭН для всех пользователей.

Вычисление средней результативности элемента представляет собой более сложную задачу, частично решаемую следующим образом. Введение таймеров на формах позволяет определить для каждого элемента, насколько тот подходит для корректного ввода данных в условиях ограниченного времени. При этом возможны три следующих качественно разных исхода тестирования:

1) пользователь уложился в имеющееся время и ввел корректно данные;

2) пользователь уложился в имеющееся время, но данные были введены некорректно;

3) пользователь не уложился в имеющееся время.

По результатам этих исходов для выборки пользователей можно будет построить график доли верно введенных данных в зависимости от времени таймера, пример которого представлен на рисунке 8.

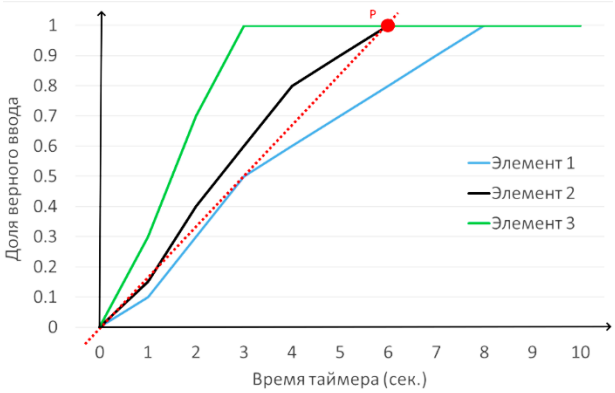


Рис. 8. Пример зависимости доли верно введенных данных от времени таймера

Fig. 8. Example of the Proportion of Correctly Entered Data on the Timer Time Dependence

Согласно графику (см. рисунок 8), производилась оценка результативности трех элементов; при этом для времени таймера более 8 секунд все они давали возможность пользователям вводить данные корректно – значение по оси ординаты равно 1. При уменьшении таймера вначале стала уменьшаться доля верно введенных данных для Элемента 1 (начиная с 8-й секунды), затем для Элемента 2 (начиная с 6-й секунды) и Элемента 3 (начиная с 3-й секунды). Таким образом, результативность Элемента 1 может считаться выше результативности Элемента 2, а затем и Элемента 3, что определяется статистической возможностью их использования для безошибочного ввода данных. Конкретные же значения результативности (естественно, в промежутке [0…1]) могут быть получены, исходя из угла наклона прямой на графике, проходящей через две точки – начало отсчета и пересечение графика с горизонтальной прямой для доли верного ответа, равного максимальной (т. е. одному); такая касательная на рисунке 8 для Элемента 2 проведена красным пунктиром, а сама точка пересечения указана как «*P*». Отметим, что другим вариантом вычисления может быть линейный тренд, также выходящий из точки «(0,0)», но учитывающий (а точнее, усредняющий) все имеющиеся значения, по которым был построен график; или же, могут использоваться более сложные аппроксимации (например, полиномами высоких степеней). Такая интерпретация результативности элемента в некотором смысле определяется максимальной скоростью работы с ним, при котором всегда обеспечивается корректный ввод.

Итоговая формула вычисления результативности имеет следующий вид:

$$\begin{cases} Potency^{Type,Length} = \dfrac{1}{Px^{Type,Length}} \\ Px^{Type,Length} = \dfrac{\sum_i Px_i^{Type,Length}}{N} \end{cases},$$

где $Px^{Type,Length}$ – проекция координаты точки «*P*» для графика (см. рисунок 8), построенного по результатам выполнения действий пользователя по вводу данных типа $Type$ и длины $Length$; $Px_i^{Type,Length}$ – аналогичная проекция точки для ввода *i*-м пользователем.

Таким образом, показатель результативности представляет собой некоторую меру того, насколько элемент позволяет корректно вводить данные при ограниченном времени. Суть данного расчета можно пояснить на двух следующих пограничных случаях. Во-первых, если элемент является условно «идеальным» (например, единственная кнопка на всю форму), позволяющим корректно вводить данные даже за минимально возможное время, то его график будет представлять горизонтальный луч из точки 1 по оси ординаты. Таким образом, проекция $Px^{Type,Length}$ будет равна 0, а сама результативность $Potency^{Type,Length} = 1$ (т. е. максимально возможной). Во-вторых, для противоположного случая, если элемент является «абсолютно неидеальным» (например, генератор случайного текста), т. е. в принципе всегда приводящим к неверному вводу, то его график будет представлять горизонтальный луч, выходящий из точки 0 по оси ординаты. Таким образом, проекция $Px^{Type,Length}$ будет стремиться к бесконечности, а сама результативность $Potency^{Type,Length}$ будет близка к 0 (т. е. минимально возможной).

Результаты

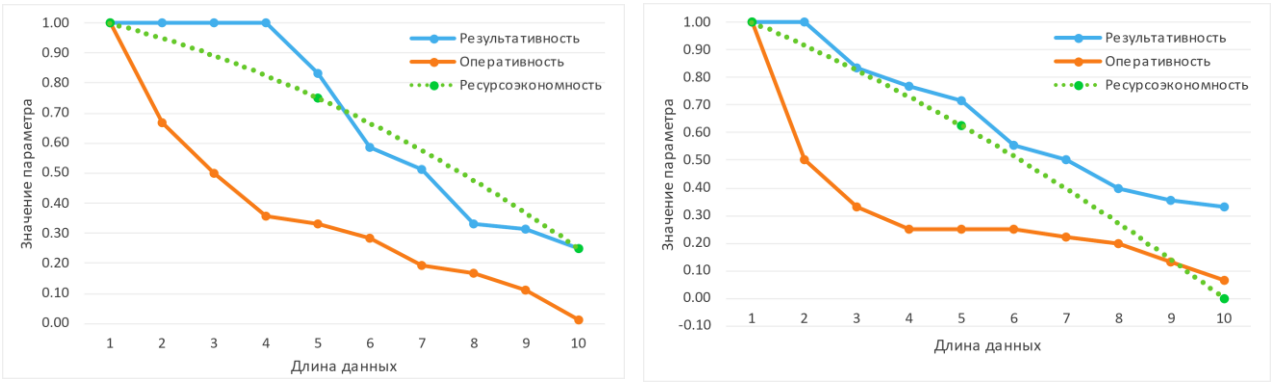
В результате применения Системой для выборки из 50 пользователей были собраны статистические данные, необходимые для определения атомарных эффективностей графических элементов. Итоговые значения для текстового поля и выпадающего списка, используемых при вводе текстовых строк различной длины (от 1 до 10 символов), полученные согласно приведенным ранее формулам, представлены на рисунке 9. Оценка ПЭН осуществлялась субъективно пользователем через соответствующую форму Системы, открывающуюся после ввода данных длиной 1, 5 и 10 символов. Соответственно, график показателя ресурсоэкономности строился по трем точкам и поэтому указан с помощью аппроксимации (полиномом второй степени) пунктирной линией; в частности, из-за этого график для выпадающего списка (см. рисунок 9b) выходит в область отрицательных чисел, что практического смысла не имеет, а обосновывается погрешностью аппроксимации.

Проведем анализ графиков зависимости атомарных эффективностей от размера вводимых данных (см. рисунок 9). Во-первых, как и следовало ожидать, результативность ввода в текстовое поле превышает аналогичный показатель для выпадаю-

щего списка только до длины ввода в 7 символов, затем же ситуация меняется на противоположную. Такую закономерность можно объяснить тем, что небольшие текстовые строки безошибочно вводятся с клавиатуры (естественно, в условиях ограничения времени таймером), а затем – данная корректность начинает обеспечиваться выбором конкретных строк из списка.

Во-вторых, оперативность ввода в текстовое поле опережает аналогичный показатель для выпадающего списка при малой длине данных – до 6 символов; затем ситуация также меняется на противоположную. Объяснение этому аналогично соотношению показателей результативности этих элементов – вводить быстрее короткие слова, а выбрать из списка – длинные строки.

В-третьих, ПЭН для двух рассмотренных элементов при небольших размерах данных – практически одинаковые (примерно для 1–2 символов), затем ПЭН для выпадающего списка начинает существенно увеличиваться – показатель ресурсоэкономности у тактового поля оказывается выше. Объяснение этого также закономерно и связано с необходимостью пользователя в случае выпадающего списка выполнять ряд дополнительных более «тяжелых» действий – пролистывание списка и поиск нужных данных.



a) b)

Рис. 9. Графики атомарных эффективностей графических элементов интерфейсов для текстового поля (a) и выпадающего списка (b)

Fig. 9. Graphs of the Graphical Interface Elements Atomic Efficiencies for Text Field (a) and Drop-Down List (b)

Заключение

Работа представляет собой один из этапов исследования автора и заключается в описании и применении разработанной Системы для статистической оценки эффективностей различных интерфейсных элементов. Принцип такого измерения состоит в определении количества ошибок, времени работы и психоэмоционального напряжения пользователя в процессе выполнения ряда заданий по вводу данных с помощью Web-интерфейса. Для снижения влияния субъективного фактора на результаты применяется ряд приемов, один из которых состоит в наличии таймера времени.

Основным результатом представленной работы является как сама Система, включающая архитектуру и способы перевода статистически измеренных характеристик работы пользователей в абсолютные значения показателей эффективности, так и конкретные зависимости последних от размера вводимых данных.

На данный момент решения, подобные предложенному, в которых бы удалось не только определить само понятие эффективности элементов интерфейса, но и предложить в достаточной степени формальный (т. е. без участия экспертов) способ ее вычисления, являются достаточно редкими. Впрочем, значение ресурсоэкономности (как ПЭН) получается на основании субъективных ощущений тестируемого, а переход в данном случае на более объективные способы оценки требует проведения некоторых дополнительных научных изысканий.

Также можно предположить значимость полученных результатов для достаточно широкого спектра задач иного рода. Например, Система способна определять предельные значения эффективности существующих и разрабатываемых графических элементов для разных режимов работы с ними, что в конечно итоге позволит создавать качественно новые решения в предметной области. Тестирование же пользователей позволит отбирать среди них наиболее подходящих для работы с интерфейсом конкретной ИС.

Продолжением работы будет следующий этап исследования (согласно разработанной ранее методологии [16]), заключающийся в создании метода оптимизации интерфейса, использующего разработанную на предыдущих этапах эффективностную модель и полученные атомарные эффективности.

Список источников

1. Мадаев С.М., Алихаджиев С.Х. Хронология развития интерфейсов, сравнение скорости выполнения задач на разных интерфейсах // Тенденции развития науки и образования. 2024. № 110-17. С. 141–144. DOI:10.18411/trnio-06-2024-952. EDN:EMOIHW
2. Доброквашина А.С. К вопросу разработки графических интерфейсов для управления БЛА // Вестник РГГУ. Серия: Информатика. Информационная безопасность. Математика. 2024. № 1. С. 8–20. DOI:10.28995/2686-679X-2024-1-8-20. EDN:ICGDBU
3. Буйневич М.В., Покусов В.В., Израилов К.Е. Способ визуализации модулей системы обеспечения информационной безопасности // Научно-аналитический журнал «Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России». 2018. № 3. С. 81–91. EDN:YKWABF
4. Курта П.А. Взаимодействие пользователя с информационной системой. Часть 1. Схема взаимодействия и классификация недостатков // Известия СПбГЭТУ ЛЭТИ. 2020. № 8-9. С. 35–45. EDN:VLVMXL
5. Курта П.А., Буйневич М.В. Онтологическая модель взаимодействия пользователя с информационной системой в рамках получения услуги информационного сервиса // Вестник кибернетики. 2021. № 2(42). С. 17–23. DOI:10.34822/1999-7604-2021-2-17-23. EDN:HSVLMI
6. Логвинов Ю.И., Горбунова Е.А., Карпова Е.В. Снижение психоэмоционального напряжения и использованием авторской методики тор (техника оптимальной ресоциализации) // Виртуальные технологии в медицине. 2020. № 1(23). С. 33–35. DOI:10.46594/2687-0037_2020_1_33. EDN:DTSOPS
7. Курта П.А. Эффективностная модель интерфейса взаимодействия пользователя с информационным сервисом запросного типа // Труды учебных заведений связи. 2023. Т. 9. № 6. С. 102–115. DOI:10.31854/1813-324X-2023-9-6-102-115. EDN:NAEEUM
8. Горина Е.В., Чебыкин К.А. Влияние элементов пользовательского интерфейса на эффективность Web-ресурса // Вестник молодых ученых Санкт-Петербургского государственного университета технологии и дизайна. 2021. № 3. С. 57–61. EDN:WBLVOR
9. 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. EDN:AIOFPM
10. Буйневич М.В., Вострых А.В. Методы оценки графических пользовательских интерфейсов. Визуальная составляющая. СПб.: Санкт-Петербургский университет Государственной противопожарной службы Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий имени Героя Российской Федерации генерала армии Е.Н. Зиничева, 2024. 340 с. EDN:EAFJIG
11. Вострых А.В. Алгоритм оценки эффективности эргономической эстетики графических пользовательских интерфейсов // Известия СПбГЭТУ ЛЭТИ. 2024. Т. 17. № 7. С. 51–61. DOI:10.32603/2071-8985-2024-17-7-51-61. EDN:NHQZFM
12. Морозов А.Н., Зарубин В.С., Гришин С.А. К вопросу оценки качества пользовательского интерфейса АРМ пунктов централизованной охраны // Вестник Воронежского института МВД России. 2019. № 1. С. 45–50. EDN:ZAQCKT
13. Каднова А.М. К вопросу оценки удобства пользовательских интерфейсов программных систем защиты информации // Общественная безопасность, законность и правопорядок в III тысячелетии. 2020. № 6-2. С. 232–235. EDN:OBCCGG
14. Кочкин А.А., Калашников С.Н., Красноперов С.Ю. Когнитивно-параметрическая оценка информационного потока пользовательского интерфейса // Электронный научный журнал. 2015. № 3(3). С. 67–72. DOI:10.18534/enj.2015.03.67. EDN:VKHDOJ
15. Shengyuan Y., Xiaoyang Y., Hongguo Z., Zhijian Z., Minjun P., Shanling W. Research of Software User Interface Evaluation Method based on Subjective Expectation // Proceedings of the International Conference on Mechatronics and Automation (Harbin, China, 05–08 August 2007). IEEE, 2007. PP. 3190–3195. DOI:10.1109/ICMA.2007.4304072
16. Буйневич М.В., Курта П.А. Методология исследования метода оптимизации информационного взаимодействия в аспекте решения задачи пользователя // Информационные технологии и телекоммуникации. 2019. Т. 7. № 4. С. 50–58. DOI:10.31854/2307-1303-2019-7-4-50-58. EDN:LULASM

References

1. Madayev S.M., Alikhadzhiev S.H. Chronology of Interfaces Development, Comparison of Task Execution Speed on Different Interfaces. *Tendencii razvitiya nauki i obrazovaniya*. 2024;110-17:141–144. (in Russ.) DOI:10.18411/trnio-06-2024-952. EDN:EMOIHW
2. Dobrokvashina A.S. Development and Testing of the Graphical User Interfaces for UAV. *RSUH/RGGU Bulletin. "Information Science. Information Security. Mathematics"*. 2024;1:8–20. (in Russ.) DOI:10.28995/2686-679X-2024-1-8-20. EDN:ICGDBU
3. Buinevich M.V., Pokusov V.V., Izrailov K.E. Method of Visualizing the Modules of the Information Security System. *Bulletin of the St. Petersburg University of the State Fire Service of the Ministry of Emergency Situations of Russia*. 2018;3:81–91. (in Russ.) EDN:YKWABF
4. Kurta P.A. Interaction of the User with the Information System. Part 1. Scheme of Interaction and Classification of Disadvantages. *Proceedings of Saint Petersburg Electrotechnical University*. 2020;8-9:35–45. (in Russ.) EDN:VLVMXL
5. Kurta P.A., Buinevich M.V. Ontological Model of User Interaction for the Purpose of Receiving Information Service. *Proceedings in Cybernetics*. 2021;2(42):17–23. (in Russ.) DOI:10.34822/1999-7604-2021-2-17-23. EDN:HSVLMI

6. Logvinov Yu.I., Gorbunova E.A., Karpova E.V. Reduce Emotional Stress and the Use of the Author's Technique Tor (The Optimal Technique of Re-Socialization). *Virtual Technologies in Medicine*. 2020;1(23):33–35. (in Russ.) DOI: 10.46594/2687-0037_2020_1_33. EDN:DTSOPS

7. Kurta P. An Efficient Interface Model of User Interaction with a Query-Type Information Service. *Proceedings of Telecommunication Universities*. 2023;9(6):102–115. (in Russ.) DOI:10.31854/1813-324X-2023-9-6-102-115. EDN:NAEEUM

8. Gorina E.V., Chebykin K.A. Influence of User Interface Elements on the Efficiency of the Web-Resource. *Vestnik molodyh uchenyh Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizajna*. 2021;3:57–61. (in Russ.) EDN:WBLVOR

9. Izrailov K. The Genetic Decompilation Concept of the Telecommunication Devices Machine Code. *Proceedings of Telecommunication Universities*. 2021;7(4):95–109. (in Russ.) DOI:10.31854/1813-324X-2021-7-4-95-109. EDN:AIOFPM

10. Buinevich M.V., Vostrykh A.V. *Methods for Evaluating Graphical User Interfaces. Visual Component*. St. Petersburg: Saint Petersburg University of State Fire Service of Emercom of Russia Publ.; 2024. 340 p. (in Russ.) EDN:EAFJIG

11. Vostrykh A.V. Algorithm for Evaluating the Effectiveness of the Ergonomic Aesthetics of Graphical User Interfaces. *Proceedings of Saint Petersburg Electrotechnical University*. 2024;17(7):51–61. (in Russ.) DOI:10.32603/2071-8985-2024-17-7-51-61. EDN:NHQZFM

12. Morozov A.N., Zarubin V.S., Grishin S.A. To the Question of Assessing the Usability of the Software Security System. *The bulletin of Voronezh Institute of the Ministry of Internal Affairs of Russia*. 2019;1:45–50. (in Russ.) EDN:ZAQCKT

13. Kadnova A.M. To the question of assessing the convenience of customer interfaces of information protection software systems. *Public Safety, Law and Order in the third millennium*. 2020;6-2:232–235. (in Russ.) EDN:OBCCGG

14. Kochkin A.A., Kalashnikov S.N., Krasnoperov S.I. Cognitive-parametric evaluation of the information flow of the user interface. *Elektronnyj nauchnyj zhurnal*. 2015;3(3):67–72. (in Russ.) DOI:10.18534/enj.2015.03.67. EDN:VKHDOJ

15. Shengyuan Y., Xiaoyang Y., Hongguo Z., Zhijian Z., Minjun P., Shanling W. Research of Software User Interface Evaluation Method based on Subjective Expectation. *Proceedings of the International Conference on Mechatronics and Automation, 05-08 August 2007, Harbin, China*. IEEE; 2007. p.3190–3195. DOI:10.1109/ICMA.2007.4304072

16. Buinevich M.V., Kurta P.A. Research Methodology of the Method of Optimization of Information Interaction in the Aspect of Solving the User's Problem. *Telecom IT*. 2019;7(4):50–58. (in Russ.) DOI:10.31854/2307-1303-2019-7-4-50-58. EDN:LULASM

Статья поступила в редакцию 18.11.2024; одобрена после рецензирования 09.12.2024; принята к публикации 18.12.2024.
The article was submitted 18.11.2024; approved after reviewing 09.12.2024; accepted for publication 18.12.2024.

Информация об авторе:

КУРТА Павел Андреевич
соискатель кафедры Прикладной математики и информационных технологий Санкт-Петербургского университета ГПС МЧС России
https://orcid.org/0009-0005-6073-8626

Автор сообщает об отсутствии конфликтов интересов.
The author declares no conflicts of interests.

Научная статья
УДК 004.732.056
https://doi.org/10.31854/1813-324X-2024-10-6-111-120

Снижение размерности массивов данных с помощью многослойных автокодировщиков в задаче классификации мобильных приложений

Олег Иванович Шелухин, sheluhin@mail.ru
Фёдор Андреевич Маторин ✉, f.matorin@mail.ru

Московский технический университет связи и информатики,
Москва, 111024, Российская Федерация

Аннотация

Рассматривается задача уменьшения размерности исходных массивов данных для улучшения эффективности обработки трафика мобильных приложений. ***Актуальность*** *исследования обусловлена необходимостью оптимизации объемов передаваемых и хранимых данных при работе в условиях ограниченных вычислительных ресурсов, а также повышения скорости и качества аналитических операций. Для решения поставленной задачи применяются многослойные автокодировщики, способные формировать компактные представления исходных данных с минимальными потерями в их информативности. Подход базируется на идее обучения нейросетевых моделей, извлекающих наиболее существенные признаки из исходных массивов и способных восстанавливать их с заданным уровнем точности.*

Используемые методы. *В ходе экспериментов применялись различные архитектуры многослойных автокодировщиков, отличающиеся количеством слоев и размерностями скрытых представлений. Исследования проводились на реальных наборах данных, собранных из мобильных приложений широкого спектра функционала. Анализ осуществлялся путем варьирования внутренних параметров сетей и оценки результатов через интегральный статистический показатель, отражающий степень сжатия. Данный показатель позволяет выявить, насколько сильно изменяется разброс атрибутов при пропускании данных через автокодировщик.*

Результаты. *Для оценки фильтрующих свойств многослойных автокодировщиков предложен интегральный показатель сжатия, характеризующий изменение разброса атрибутов мобильных приложений при пропускании их через автокодировщик заданной структуры. Показатель рассчитывается как отношение среднеквадратического отклонения атрибутов на входе и на выходе, что позволяет оценить степень сжатия данных и степень сохранности информации после обработки. Показано, что увеличение интегрального показателя сжатия свидетельствует о более значительном сжатии исходных данных. Установлено, что фильтрация практически не зависит от типа приложения и лежит в пределах 10–20 % для автокодировщиков с тремя слоями, тогда как для пятислойных автокодировщиков предпочтение отдается кодировщикам с минимальной размерностью внутреннего слоя. Основная* ***новизна*** *работы заключается в разработке интегрального статистического показателя, который не только отражает степень сжатия данных мобильных приложений, но и учитывает сохранность исходной информационной структуры. В отличие от существующих подходов, данный показатель позволяет проводить систематическое сравнение различных архитектур автокодировщиков с учетом не только уменьшения размерности, но и качества восстановления исходной информации. Это создает основу для более объективной оценки эффективности многослойных автокодировщиков в конкретных прикладных условиях.* ***Практическая значимость****. Предложенная методология может быть полезна разработчикам и исследователям, работающим над оптимизацией систем сбора, хранения и обработки данных мобильных приложений. В условиях ограниченных вычислительных ресурсов, характерных для мобильных устройств и встроенных систем, использование многослойных автокодировщиков, настроенных на достижение заданного баланса между сжатием и сохранением информации, обеспечивает существенное сокращение объема передаваемых данных. Результаты исследования могут быть внедрены в существующие аналитические платформы, системы мониторинга и классификации мобильных приложений.*

Ключевые слова: *нейронные сети, классификация, приложения, атрибуты, фильтрация, статистические характеристики*

Ссылка для цитирования: Шелухин О.И., Маторин Ф.А. Снижение размерности массивов данных с помощью многослойных автокодировщиков в задаче классификации мобильных приложений // Труды учебных заведений связи. 2024. Т. 10. № 6. С. 111–120. DOI:10.31854/1813-324X-2024-10-6-111-120. EDN:TOPDUA

Original research
https://doi.org/10.31854/1813-324X-2024-10-6-111-120

Reducing the Dimensionality of Data Arrays Using Multi-Layer Autoencoders in the Task of Classifying Mobile Applications

Oleg I. Sheluhin, sheluhin@mail.ru
Fedor A. Matorin✉, f.matorin@mail.ru

Moscow Technical University of Communications and Informatics,
Moscow, 111024, Russian Federation

Annotation

The problem of reducing the dimension of the initial data arrays to improve the efficiency of mobile application traffic processing is considered. ***The relevance*** *of the study is due to the need to optimize the volume of transmitted and stored data when working in conditions of limited computing resources, as well as to increase the speed and quality of analytical operations. To solve this problem, multi-layer autoencoders are used, capable of forming compact representations of the source data with minimal losses in their informativeness. The approach is based on the idea of training neural network models that extract the most significant features from the source arrays and are able to restore them with a given level of accuracy.* ***Methods used.*** *During the experiments, various architectures of multilayer autocoders were used, differing in the number of layers and dimensions of hidden representations. The research was conducted on real data sets collected from mobile applications with a wide range of functionality. The analysis was carried out by varying the internal parameters of the networks and evaluating the results through an integral statistical indicator reflecting the degree of compression. This indicator allows you to identify how much the spread of attributes changes when passing data through the autoencoder.*

Results. *To evaluate the filtering properties of multilayer autoencoders, an integral compression indicator is proposed that characterizes the change in the spread of attributes of mobile applications when passing them through an autoencoder of a given structure. The indicator is calculated as the ratio of the standard deviation of the attributes at the input and at the output, which allows you to assess the degree of data compression and the degree of information preservation after processing. It is shown that an increase in the integral compression index indicates a more significant compression of the initial data. It was found that filtering is practically independent of the type of application and lies within 10–20 % for three-layer autoencoders, whereas for five-layer auto-encoders, preference is given to encoders with a minimum dimension of the inner layer. The main* ***novelty*** *of the work lies in the development of an integral statistical indicator that not only reflects the degree of compression of mobile application data, but also takes into account the preservation of the original information structure. Unlike existing approaches, this indicator allows for a systematic comparison of various architectures of autoencoders, taking into account not only the reduction in dimension, but also the quality of recovery of the original information. This creates the basis for a more objective assessment of the effectiveness of multilayer autoencoders in specific application conditions.*

Practical significance. *The proposed methodology may be useful for developers and researchers working on optimizing systems for collecting, storing and processing mobile application data. In conditions of limited computing resources, which are typical for mobile devices and embedded systems, the use of multilayer autoencoders aimed at achieving a given balance between compression and preservation of information provides a significant reduction in the volume of transmitted data. The results of the study can be implemented into existing analytical platforms, monitoring systems and classification of mobile applications.*

Keywords: *neural networks, classification, applications, attributes, filtering, statistical characteristics*

For citation: Sheluhin O.I., Matorin F.A. Reducing the Dimensionality of Data Arrays Using Multi-Layer Autoencoders in the Task of Classifying Mobile Applications. *Proceedings of Telecommunication Universities.* 2024;10(6):111–120. (in Russ.) DOI:10.31854/1813-324X-2024-10-6-111-120. EDN:TOPDUA

Постановка задачи

Снижение размерности данных играет ключевую роль в задачах анализа больших массивов информации, особенно в контексте обработки данных мобильных приложений. Эти приложения генерируют значительные объемы сетевого трафика, который часто содержит избыточную информацию. Эффективная фильтрация и сжатие таких данных позволяют уменьшить объем обрабатываемой информации, снижая нагрузку на сеть и требования к вычислительным ресурсам. Это особенно актуально в условиях ограниченной полосы пропускания и низких мощностей мобильных устройств. Для снижения размерности данных мо-

гут использоваться различные методы, основанные на искусственных нейронных сетях.

Глубокие сверточные нейронные сети [1], эффективно обрабатывающие данные с пространственной структурой, что делает их отличным выбором для работы с изображениями. Однако использование глубоких сверточных нейронных сетей требует значительного объема размеченных данных, что не всегда возможно при работе с неструктурированными данными, такими как сетевой трафик мобильных приложений.

Сети глубокого доверия [2], позволяющие поэтапно обучать модель, снижая размерность данных. Однако подобные сети требуют тщательной настройки большого количества гиперпараметров, что может затруднять их использование в условиях ограниченных вычислительных ресурсов [3].

Ограниченные машины Больцмана [4] являются эффективным инструментом для выявления скрытых зависимостей в данных, однако их обучение является вычислительно затратным при увеличении размерности данных и сложности структуры.

Многослойные автокодировщики [1, 5–7] позволяют использовать неконтролируемое обучение, что делает их особенно подходящими для работы с неразмеченными данными. Они обеспечивают баланс между эффективностью сжатия и сохранением значимой информации, что особенно важно для задач обработки мобильных приложений.

Опираясь на рассмотренные методы и работы, можно сделать вывод, что многослойные автокодировщики (АК) являются наиболее подходящим выбором для задач, связанных с обработкой трафика мобильных приложений, благодаря их способности к неконтролируемому обучению и высокой эффективности при работе с неразмеченными данными. В работах [8, 9] была доказана эффективность применения АК в задачах классификации нежелательных мобильных приложений, что подтверждает их целесообразность использования в данной области. Выбор многослойных АК для обработки данных мобильных приложений позволяет найти баланс между сжатием и сохранением значимой информации, а также позволяют работать с неразмеченными данными, что снижает требования к предварительной подготовке данных и улучшает эффективность обработки.

Целью работы является исследование влияния многослойных АК на фильтрацию и снижение размерности обрабатываемых данных мобильных приложений с целью улучшения эффективности их классификации и обработки. Достижение этой цели позволит сократить объем данных, которые необходимо передавать и хранить, что приведет к снижению нагрузки на вычислительные ресурсы и сети передачи данных, а также увеличит точность классификации приложений, что особенно важно в задачах обеспечения кибербезопасности, автоматического контроля и оптимизации работы мобильных сервисов.

Модели многослойных автокодировщиков

Основой для построения всех моделей многослойных АК является модель простого трехслойного АК. Это сеть прямого распространения с входным и выходным слоями, содержащими одинаковое число нейронов, и единственным внутренним (горловым) слоем, содержащим меньшее число нейронов, чем входной и выходной слои.

Будем считать $X_1, X_2, \dots X_M \in R^N$ векторами входных данных, характеризующими анализируемые мобильные приложения. Тогда матрицу входных данных можно представить в виде:

$$X = [X_1, X_2, \dots X_M]^T,$$

где каждая строка представляет собой вектор обрабатываемых признаков (атрибутов) M анализируемых приложений, а число столбцов N характеризует размерность пространства признаков.

В результате матрица X представляет собой матрицу размера $N \times M$:

$$X = \begin{vmatrix} X_{11} & X_{12} & \cdots & X_{1M} \\ X_{21} & X_{22} & \cdots & X_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N1} & X_{N2} & \cdots & X_{NM} \end{vmatrix}, X \in R^{N \times M}.$$

Рассмотрим структуру предназначенного для сокращения размерности больших массивов данных, подлежащих обработке, многослойного АК [4, 10], который представляет собой специальный вид сети прямого распространения (МАК-сеть) – многослойный симметричный персептрон, содержащий несколько внутренних слоев уменьшающегося размера и слой «бутылочная горловина» в середине сети. МАК-сеть производит тождественное преобразование входного слоя на выходной. В результате ее работы в горловом слое появляется вектор, компонентами которого являются «признаки» – обобщенные характеристики входного массива данных, извлеченные из исходных данных и содержащие дополнительную существенную и не избыточную информацию, определяющую входной массив данных в пространстве меньшей размерности в так называемом скрытом пространстве.

Задачей скрытого пространства является выделение важных признаков (атрибутов), которые будут использоваться для восстановления исходных данных при максимально малой размерности слоя. Структура простейшего трехслойного АК представлена на рисунке 1.

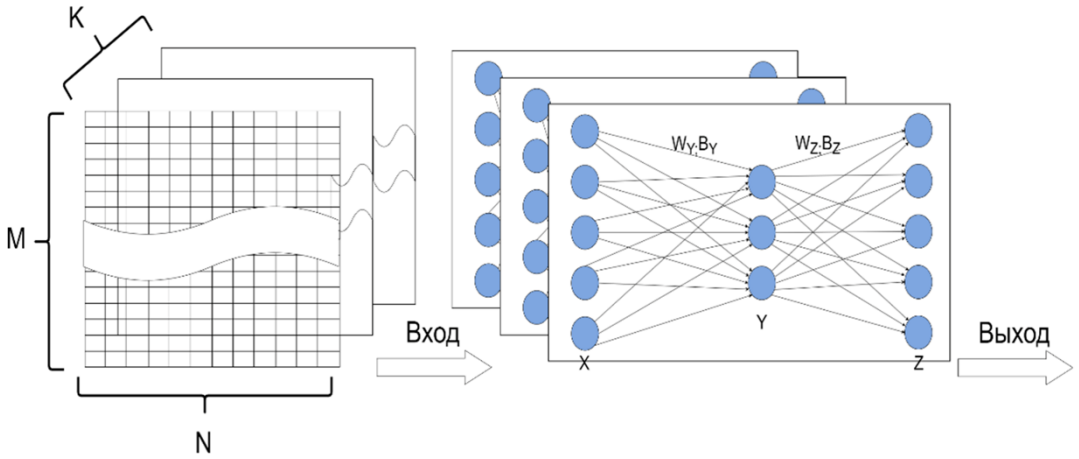


Рис. 1. Структура трехслойного АК

Fig. 1. Structure of a Three-Layer Autoencoder (AE)

Как показано на рисунке 1, наиболее простой АК представляет собой многослойный персептрон, который имеет один скрытый слой и один выходной с двумя ограничениями: матрица весов выходного слоя является транспонированной матрицей весов скрытого слоя $\widehat{W}_Y = \widehat{W}_Z^T = \widehat{W}$ (т. е. веса фиксированы) и количество выходных нейронов равно количеству входных.

Значения нейронов скрытого слоя, называемые кодированием, вычисляются по выражению:

$$Y = G_\theta(X) = F\left(\widehat{W}_Y X + B_Y\right),\ \theta = \{W_Y, B_Y\}, \qquad (1)$$

где X – входной вектор; F – функция активации нейронов сети; B_Y – вектор скрытых нейронных смещений; W_Y – матрица скрытых весов.

Задача функции кодирования $Y = F(X, \widehat{W}_Y, B_Y)$ заключается в сжатии входного вектора в соответствии с уравнениями.

Операция декодирования характеризуется функцией декодирования $Z = F(Y, \widehat{W}_Z, B_Z)$ и заключается в восстановлении входного «сжатого» вектора:

$$Z = G_{\hat{\theta}}(Y) = F\left(\widehat{W}_Z Y + B_Z\right),\ \hat{\theta} = \{W_Z, B_Z\}. \qquad (2)$$

В формулах (1) и (2) $\widehat{W}_Y$ и $\widehat{W}_Z$ – матрицы сетевых связей (матрицы весовых коэффициентов) кодировщика и декодировщика АК; весовые коэффициенты B_Y и B_Z – векторы смещения (определяют важность каждого входного сигнала для вычисления выходных значений слоя); θ и $\hat{\theta}$ – наборы параметров отображения.

Каждый нейрон имеет свое собственное смещение, не зависящее от входных данных, и настраивается в процессе обучения модели вместе с весами. Количество весов определяется количеством нейронов на предыдущем слое, а количество смещений – количеством нейронов на текущем. Общее количество параметров определяется соотношением:

$$(input_{demention} + 1) * cur_{dence_{demention}},$$

где $input_{demention}$, $cur_{dence_{demention}}$ – размерность предыдущего (размер выборки) и текущего слоя, соответственно, или значение веса и смещения.

Размерности скрытых слоев зависят от желаемой степени сжатия входных данных, количества признаков выборки и целевого значения размерности скрытого пространства – параметра, который влияет на способность модели к обучению и реконструкции. Слой Y содержит меньшее количество информативных параметров обрабатываемого массива данных, извлеченных в процессе работы АК. Меньшая размерность скрытого слоя может привести к более эффективному сжатию и выделению значимых признаков. Однако при этом увеличивается риск потери информации, и наоборот.

Целью обучения АК является минимизация разницы между входными X и выходными Z данными.

Типичная функция потерь представляет собой среднеквадратическую ошибку (СКО):

$$L(X, Y) = \|X - Z\|^2. \qquad (3)$$

Используя (1) и (2), выражение (3) может быть преобразовано к следующему виду:

$$L(X, Y) = \left\| X - F\left(\widehat{W}_Z\left(F\left(\widehat{W}_Y X + B_Y\right)\right) + B_Z\right)\right\|^2. \qquad (4)$$

Функция потерь $L(X, Y)$ определяет качество реконструкции оригинала, так что выходная реконструкция должна быть как можно ближе к исходному входному вектору. Отсюда основной задачей является минимизация значений функции потерь и обновление ее параметров для повышения точности реконструкции.

Наиболее распространенными функциями потерь являются СКО (MSE) и корень из СКО (RMSE).

Настройка многослойных АК осуществляется путем минимизации функции потерь:

$$L(X,Y) = cost(X,Y),$$

которая может быть произведена различными способами, например методом градиентного спуска, и позволяет обновлять параметры для повышения точности.

Основная цель обучения многослойных АК состоит в том, чтобы найти оптимальные параметры (θ и $\hat{\theta}$), которые могут эффективно минимизировать разницу между входными и восстановленными выходными данными по всему обучающему набору:

$$\theta = \{W, B\} = \arg_{\theta}\min L(X,Y). \quad (5)$$

Работу многослойной нейронной сети прямого распространения можно интерпретировать как вычисление композиции многомерных отображений многослойного АК, содержащего несколько внутренних слоев, которые обладают большими возможностями по сравнению с простыми трехслойными АК. В качестве примера на рисунке 2 изображена структура пятислойного АК.

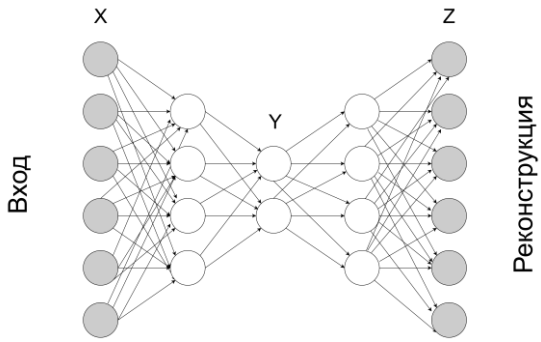


Рис. 2. Структура пятислойного АК

Fig. 2. Structure of a Five-Layer AE

Для многослойного АК можно записать:

$$Z = G\left(F(X;\theta), \theta\left(\{W_{jk}\}, \{V_k\}\right)\right),$$

где значения θ, обеспечивающих наилучшую аппроксимацию композиции функций, находятся путем обучения АК.

Вектор состояния слоя j; $Z^{(j)} \in R^{Lj}$, преобразуется в вектор состояния слоя j+1; $Z^{(j+1)} \in R^{j+1}$:

$$Z^{(j+1)} = \widehat{W}^{(i)} F\left(Z^{(j)}\right) + B^{(j)}, \quad (6)$$

где $\widehat{W}^{(i)}$ – веса $(L_j \times L_{j+1})$ матрицы связей слоев j и j+1; F – функция активации нейронной сети.

В структуре АК могут использоваться различные функции активации нейронов сети [11]. Нелинейность функции активации позволяет извлекать из исходных данных более существенные обобщенные характеристик, устраняя как линейные, так и нелинейные корреляции.

Реализация АК подразумевает под собой конфигурирование слоев кодера и декодера, указание функций активации для них, выбор гиперпараметров и оптимизацию параметров АК. Кроме перечисленных параметров при использовании многослойных АК, в задачах классификации и прогнозирования необходимо задаться гиперпараметрами. Для моделей многослойных АК гиперпараметрами являются веса и смещения внутри каждого слоя кодировщика и декодировщика. Эти параметры определяют, как модель сжимает входные данные и восстанавливает их обратно.

Оценка фильтрующих свойств многослойных автокодировщиков

Рассмотрим оценку фильтрующих свойств многослойных АК на примере экспериментальных данных мобильных приложений, приведенных в работах [12, 13]. Для формирования обучающей и тестовой выборок на мобильных устройствах под управлением ОС Android осуществлялся сбор необработанных данных сетевого трафика в виде IP-пакетов. Обработка данных (в том числе фильтрация пакетов, содержащих данные протокола TCP, группировка пакетов в TCP-сеансы и вычисление их атрибутов, характеризующих особенности анализируемых приложений) осуществлялась на сервере всякий раз, когда поступал IP-пакет. С применением разработанного программного комплекса был собран трафик различных типов мобильных приложений, из которых в дальнейшем будем использовать M = 6 мобильных приложений (*Skype*, *Booking*, *Instagram* (Деятельность Meta Platform Inc. по реализации продуктов – социальных сетей Facebook и Instagram на территории РФ запрещена по основаниям осуществления экстремистской деятельности), *Mail*, *SberMobile*). Каждое из них описывается набором из N = 21 атрибута, характеризующих то или иное приложение. Общее число экспериментально измеренных потоков каждого приложения составляло K = 5000 измерений.

Фильтрующую способность многослойных АК будем характеризовать динамическим диапазоном изменения разброса численных значений атрибутов исследуемых приложений до и после обработки данных с помощью АК. Эффект фильтрации можно проиллюстрировать гистограммами распределения одного из атрибутов приложения *Mail* до и после АК, представленными на рисунке 3. Как видно из рисунка, специфика структуры и обработки в АК приводит к уменьшению динамического диапазона изменения численных значений атрибутов на выходе АК, что иллюстрирует эффект сжатия (фильтрации). Исследовались структуры многослойных АК с тремя и пятью слоями и сигмоидальной функцией активации.

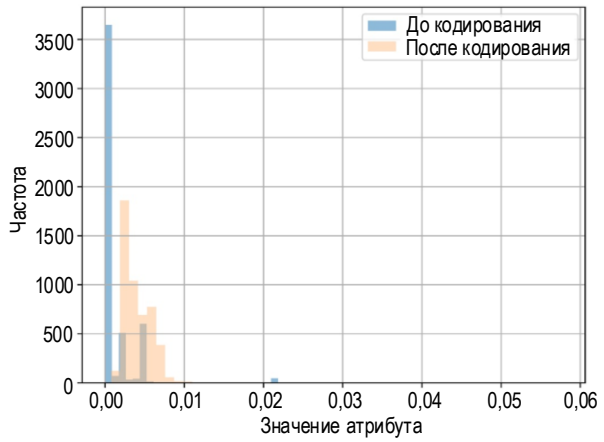


Рис. 3. Гистограмма атрибута *i* = 18 приложения *j* = 5 (*Mail*) до и после АК: $\sigma_{18,5\ вх} = 0{,}0126$; $\sigma_{18,5вых} = 0{,}0047$

Fig. 3. Histogram of Attribute i = 18 of Application j = 5 (Mail) before and after AE: $\sigma_{18,5\ вх} = 0{,}0126$; $\sigma_{18,5вых} = 0{,}0047$

Для формализации параметра, характеризующего фильтрующую способность многослойных АК введем в рассмотрение следующие обозначения:

$\sigma^2_{ij\text{вх}} = \frac{1}{K}\sum_{K=1}^{K}(a^{\text{вх}}_{ijk} - ma^{\text{вх}}_{ijk})^2$ – дисперсия *i*-го атрибута *j*-го приложения на входе многослойного АК;

$ma^{\text{вх}}_{ijk} = \frac{1}{K}\sum_{K=1}^{K} a^{\text{вх}}_{ijk}$ – среднее значение *i*-го атрибута *j*-го приложения, посчитанное по *K* измерениям;

$\sigma_{ij\text{вх}} = \sqrt{\sigma^2_{ij\text{вх}}}$ – СКО *i*-го атрибута *j*-го приложения, на входе многослойного АК;

$\sigma^2_{ij\text{вых}} = \frac{1}{K}\sum_{K=1}^{K}(a^{\text{вых}}_{ijk} - ma^{\text{вых}}_{ijk})^2$ – дисперсия *i*-го атрибута *j*-го приложения на выходе многослойного АК;

$ma^{\text{вых}}_{ijk} = \frac{1}{K}\sum_{K=1}^{K} a^{\text{вых}}_{ijk}$ – среднее значение *i*-го атрибута *j*-го приложения на выходе многослойного АК, посчитанное по *K* измерениям;

$\sigma_{ij\text{вых}} = \sqrt{\sigma^2_{ij\text{вых}}}$ – СКО *i*-го атрибута *j*-го приложения на выходе многослойного АК;

$\Delta\sigma_{ij} = \sigma_{ij\text{вх}} - \sigma_{ij\text{вых}}$ – абсолютное изменение СКО *i*-го атрибута *j*-го приложения на выходе многослойного АК по сравнению со входом;

$\delta_{ij} = \frac{\Delta\sigma_{ij}}{\sigma_{ij\text{вх}}} * 100\ \%$ – относительное изменение СКО *i*-го атрибута *j*-го приложения на выходе многослойного АК по сравнению со входом;

$\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij\text{вх}}$ – усредненное по всем атрибутам среднее значение СКО *j*-го приложения на входе многослойного АК;

$\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij\text{вых}}$ – усредненное по всем атрибутам среднее значение СКО *j*-го приложения на выходе многослойного АК;

$\frac{1}{N}\sum_{i=1}^{N}\Delta\sigma_{ij}$ – усредненное по всем атрибутам среднее значение абсолютного изменения СКО *j*-го приложения на выходе многослойного АК по сравнению со входом;

$\frac{1}{N}\sum_{i=1}^{N}\delta_{ij}\ \%$ – усредненное по всем атрибутам среднее *относительное* изменение СКО *j*-го приложения на выходе многослойного АК по сравнению со входом.

Введенную в рассмотрение величину, которую можно описать следующим выражением:

$$\text{ИПС}j = \frac{1}{N}\sum_{i=1}^{N}\delta_{ij}\ \% = \frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\sigma_{ij}}{\sigma_{ij\text{вх}}} * 100\ \% = \\ = \frac{1}{N}\sum_{i=1}^{N}\frac{\sigma_{ij\text{вх}} - \sigma_{ij\text{вых}}}{\sigma_{ij\text{вх}}} * 100\ \%, \quad (7)$$

будем называть интегральным статистическим показателем (ИПС, *от англ.* Integral Statistic – *IS*) сжатия многослойного АК *j*-го приложения и использовать его для оценки фильтрующей способности многослойного АК заданной структуры. Данная величина является интегральным показателем, характеризующим изменение разброса атрибутов рассматриваемых приложений при пропускании через многослойный АК заданной структуры. Чем больше величина ИПС, тем значительней СКО входного показателя больше СКО выходного показателя. В качестве примера в таблице 1 приведены промежуточные результаты оценки введенных метрик для многослойного АК с тремя слоями.

ТАБЛИЦА 1. Метрики для АК с 3 слоями и структурой 21-5-21

TABLE 1. Metrics for AE with 3 Layers and Structure 21-5-21

j	Приложения	$\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij\text{вх}}$	$\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij\text{вых}}$	$\frac{1}{N}\sum_{i=1}^{N}\Delta\sigma_{ij}$	ИПС, %
1	*Chrome*	0,1496	0,1406	0,0090	10,2823
2	*Yandex*	0,0948	0,0820	0,0129	30,6003
3	*Booking*	0,0915	0,0728	0,0188	38,0363
4	*ISG**	0,0993	0,0873	0,0119	21,4529
5	*Mail*	0,0898	0,0770	0,0128	32,9408
6	*SberMobile*	0,1496	0,1406	0,0090	10,2823

Значения ИПС для последовательности каждого из 21-го атрибута приложения при использовании АК с 3-мя слоями и структурой представлены на рисунке 4. Анализ представленных зависимостей показывает, что у АК с тремя слоями наблюдается выигрыш в фильтрующей способности обусловленный уменьшением СКО процесса на выходе кодировщика. Зависимости уменьшения разброса выходных данных АК от анализируемых атрибутов, представленные на рисунке 4, показывают, что выигрыш слабо зависит от типа приложения и лежит в среднем в пределах 10…20 % за исключением атрибутов № 2, 3, 13, 18, у которых выигрыш достигает 60…100 %. Сравнительный анализ эффективности АК с тремя слоями оцениваемый по-

** Деятельность Meta Platform Inc. по реализации продуктов – социальных сетей Facebook и Instagram на территории РФ запрещена из-за экстремистской деятельности*

казателем ИПС иллюстрируется гистограммами, приведенными на рисунке 5.

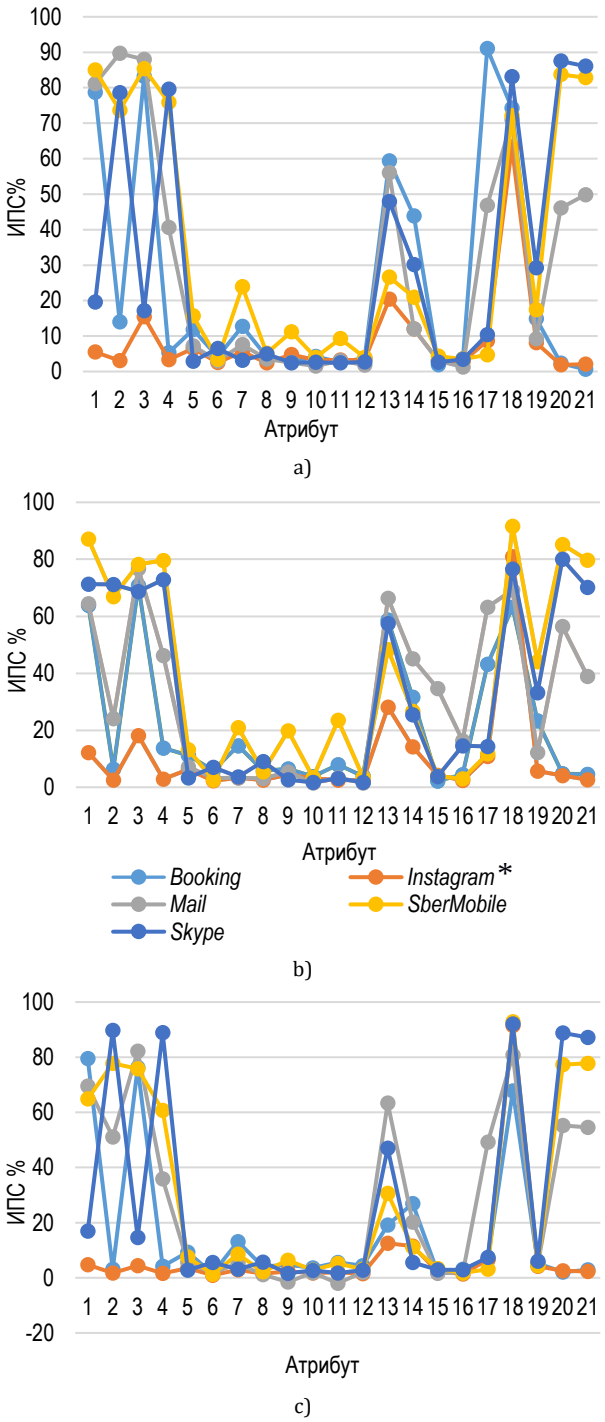


Рис. 4. Значение ИПС для последовательности атрибутов приложения, при использовании АК с 3-мя слоями и структурой: a) 21-7-21; b) 21-5-21; c) 21-9-21

Fig. 4. IS Value for a Sequence of Application Attributes, Using an AE with 3 Layers and Structure: a) 21-7-21; b) 21-5-21; c) 21-9-21

Гистограммы показывают, что в среднем уменьшение разброса обрабатываемых данных, оцениваемое величиной СКО, наиболее предпочтительно для АК с тремя слоями и структурой 21-5-21. Для этой структуры выигрыш достигает 20…25 % независимо от типа приложения.

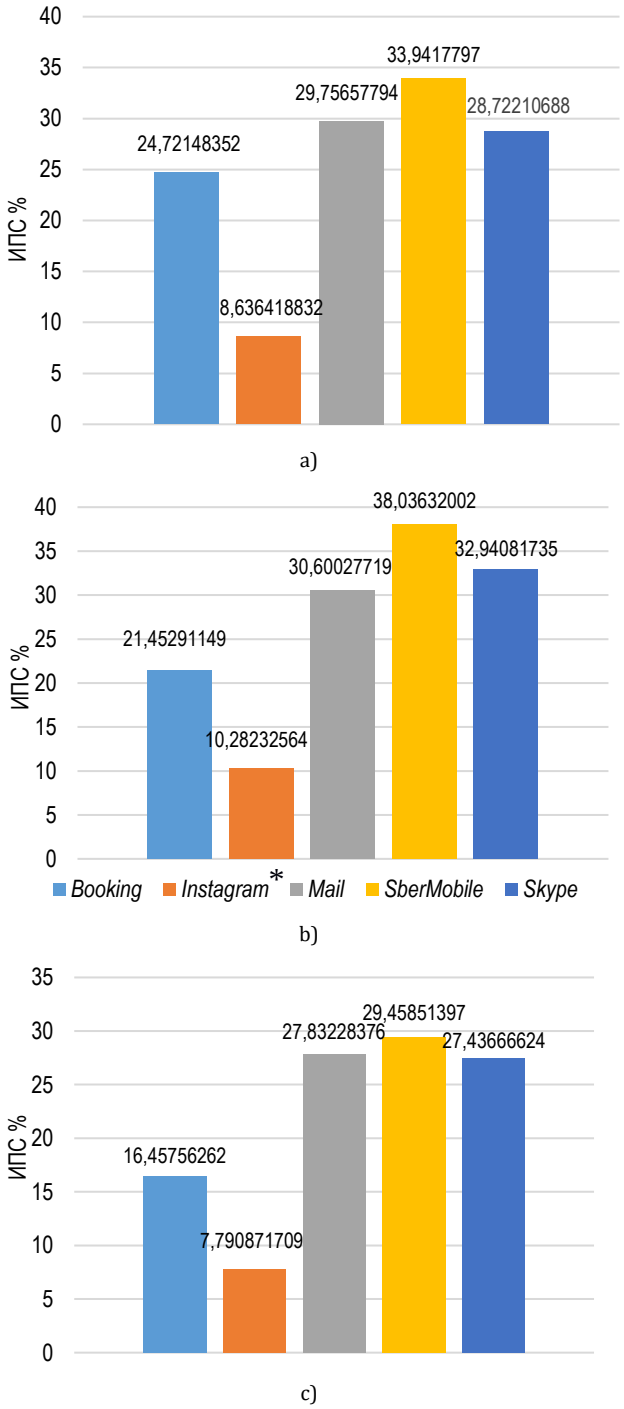


Рис. 5. Гистограммы среднего значения ИПС в процентах от приложения, при использовании АК с 3-мя слоями и структурой: a) 21-7-21; b) 21-5-21; c) 21-9-21

Fig. 5. Histograms of the Average IS Value in Percent of the Application, Using AE with 3 Layers and Structure: a) 21-7-21; b) 21-5-21; c) 21-9-21;

Для структуры 21-7-21 выигрыш скромнее и достигает в среднем 15…20 %, а для структуры 21-9-21 – не превышает 15 %.

* Деятельность Meta Platform Inc. по реализации продуктов – социальных сетей Facebook и Instagram на территории РФ запрещена из-за экстремистской деятельности

Таким образом, для анализа эффективности АК в задаче классификации нежелательных приложений целесообразно ограничиться структурой с наименьшим размером внутреннего слоя 21-5-21.

Анализ фильтрующих свойств многослойных АК для количества слоев более трех проводился для числа слоев равного 5. Численные значения для АК с 5 слоями и структурой 21-14-5-14-21 для рассмотренных выше метрик представлены в таблице 2. На рисунке 6 представлены зависимости ИПС при использовании АК с 5-ю слоями и различной структурой слоев. Анализ многослойных АК с пятью слоями и структурами 21-14-5-14-21 (см. рисунки 6c, 6d) и 21-14-7-14-21 (см. рисунки 6a, 6b) показывает, что зависимости от вида атрибутов сохраняются, как и для АК с тремя слоями. Выигрыш в среднем не превышает 10 % за исключением атрибутов с № 13, 17, в которых выигрыш может достигать 80…100 %.

ТАБЛИЦА 2. Метрики для АК с 5-ю слоями и структурой 21-14-5-14-21

TABLE 2. Metrics for AE with 5 Layers and Structure 21-14-5-14-14-21

j	Приложения	$\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij\text{вх}}$	$\frac{1}{N}\sum_{i=1}^{N}\sigma_{ij\text{вых}}$	$\frac{1}{N}\sum_{i=1}^{N}\Delta\sigma_{ij}$	ИПС, %
1	*ISG**	0,1496	0,1446	0,0050	7,6304
2	*Mail*	0,0948	0,0896	0,0052	15,7605
3	*SberMobile*	0,0915	0,0872	0,0044	11,1490
4	*Booking*	0,0993	0,0898	0,0095	19,0898
5	*Chrome*	0,0898	0,0834	0,0064	21,9972
6	*Yandex*	0,1496	0,1446	0,0050	7,6304

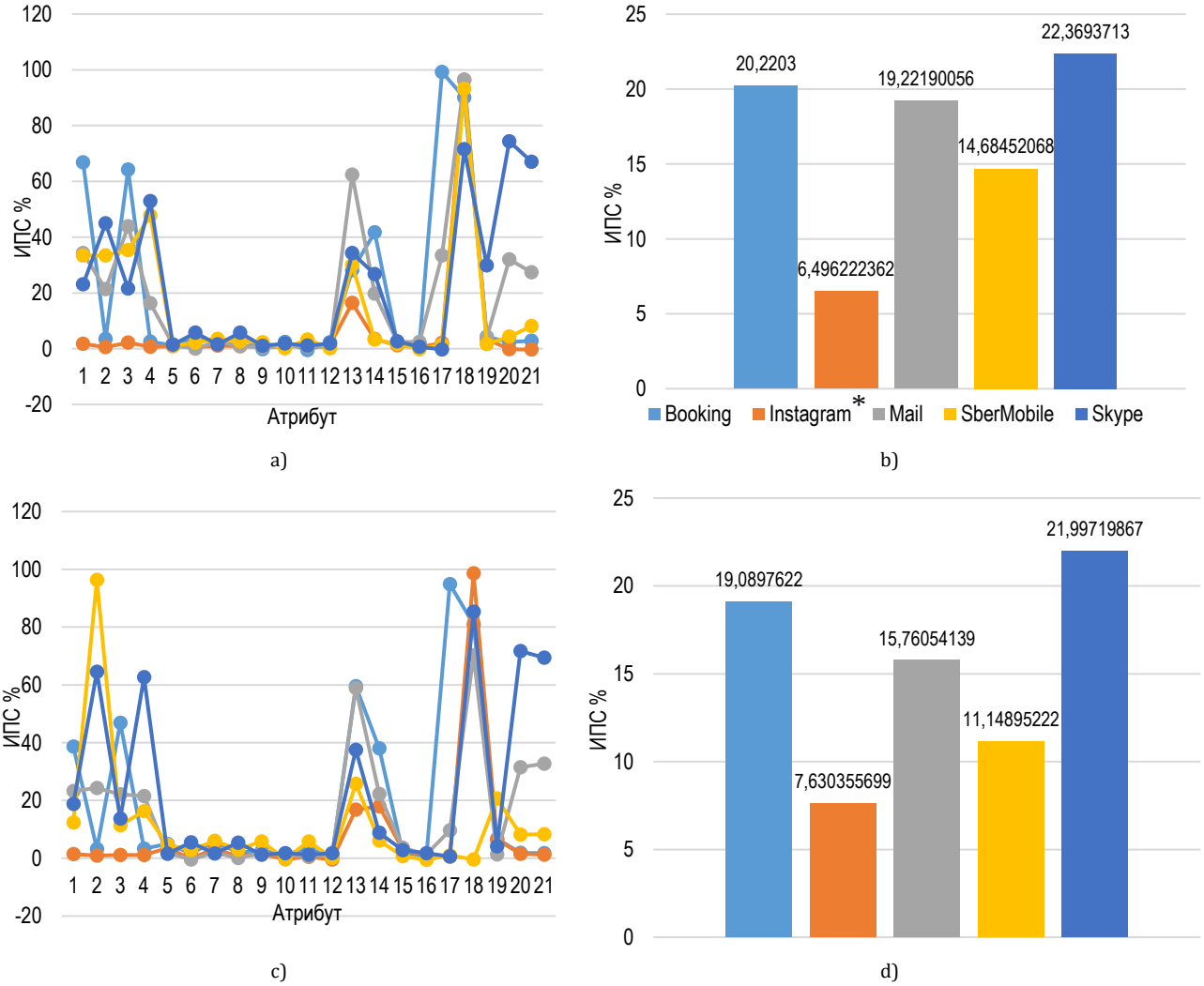


Рис. 6. Зависимости ИПС при использовании АК с 5-ю слоями от типа атрибута для рассматриваемых приложений (слева) и от гистограммы распределения ИПС для различных приложений (справа)

Fig. 6. Dependencies of IS Using AE with 5 Layers on the Attribute Type for the Considered Applications (Left) and from the Histogram of IS Distribution for Different Applications (Right)

* Деятельность Meta Platform Inc. по реализации продуктов – социальных сетей Facebook и Instagram на территории РФ запрещена из-за экстремистской деятельности

Что касается зависимости среднего ИПС от типа приложения, при использовании многослойного АК с 5-ю слоями со структурой 21-14-5-14-21, то она в среднем составляет 15 %, а для структуры 21-14-7-14-21 – около 6 %. Это показывает, что при использовании многослойных АК с 5-ю слоями предпочтение имеет кодировщик, размерность внутреннего слоя которого минимальна.

Сравнение многослойного АК с тремя слоями и структурой 21-5-21 и пятью слоями и структурой 21-14-5-14-21 показывает, что предпочтение следует отдать АК с тремя слоями, в которой выигрыш в среднем составляет 23 %.

Заключение

В работе были исследованы фильтрующие свойства многослойных автокодировщиков в задачах снижения размерности данных мобильных приложений. Основным научным результатом работы является разработка интегрального статистического показателя сжатия, который позволяет количественно оценить изменение разброса атрибутов мобильных приложений после обработки автокодировщиком заданной структуры.

Введенный показатель позволяет оценить эффективность сжатия данных и степень сохранения важной информации. Показано, что чем больше его величина, тем значительней среднеквадратическая ошибка входного показателя выше величины среднеквадратической ошибки выходного показателя, и тем лучше осуществляется сжатие входных данных. Зависимости уменьшения разброса выходных данных многослойных АК от анализируемых атрибутов (см. рисунок 1) показывают, что выигрыш слабо зависит от типа приложения и лежит в среднем в пределах 10…20 % за исключением отдельных атрибутов.

В результате экспериментов было установлено, что трехслойные автокодировщики со структурой 21-5-21 обеспечивают наилучший баланс между сжатием и сохранением информации, достигая уменьшения разброса данных на 20…25 %. Для пятислойных автокодировщиков предпочтение отдается моделям с минимальной размерностью внутреннего слоя, так как они обеспечивают наименьшие потери информации.

Применение разработанного статистического показателя дает возможность сравнивать разные конфигурации автокодировщиков не только по степени уменьшения объема данных, но и по качеству восстановления исходных признаков. Таким образом, проведенное исследование расширяет арсенал методологических средств, доступных специалистам в области анализа данных мобильных приложений, и формирует предпосылки для более осмысленного выбора архитектурных параметров нейросетевых моделей.

Практическая значимость полученных результатов проявляется в возможности адаптации разработанного подхода к реальным приложениям, где ограниченные ресурсы и требования к скорости обработки данных играют ключевую роль. Использование предложенной методологии помогает снижать затраты на хранение и передачу данных, ускорять аналитические операции и, в конечном счете, повышать общую эффективность мобильных сервисов, делая их более производительными и надежными.

Список источников

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. The MIT Press, 2016. 800 p.
2. Hinton G.E., Osindero S., Teh Y.W. A Fast Learning Algorithm for Deep Belief Nets // Neural Computation. 2006. Vol. 18. Iss. 7. PP. 1527–1554. DOI:10.1162/neco.2006.18.7.1527
3. Salakhutdinov R., Hinton G.E. Deep Boltzmann Machines // Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (Clearwater Beach, USA). Proceedings of Machine Learning Research. 2009. Vol. 5. PP. 448–455.
4. Кузьмина М.Г. Многослойные сети-автоэнкодеры в задачах анализа и обработки гиперспектральных изображений // Препринты ИПМ им. М. В. Келдыша. 2021. № 28. 21 с. DOI:10.20948/prepr-2021-28
5. Kramer M.A. Nonlinear principal component analysis using autoassociative neural networks // AIChE Journal. 1991. Vol. 37. Iss. 2. PP. 233–243. DOI:10.1002/aic.690370209
6. Bengio Y., Lamblin P., Popovici D., Larochelle H. Greedy Layer-Wise Training of Deep Networks // In: Advances in Neural Information Processing Systems (B. Schölkopf, J. Platt, T. Hoffman (eds.). Cambridge, 2007. PP. 153–160.
7. Windrim L., Ramakrishnan R., Melkumyan A., Murphy R.J., Chlingaryan A. Unsupervised feature-learning for hyperspectral data with autoencoders // Remote Sensing. 2019. Vol. 11. Iss. 7. P. 864. DOI:10.3390/rs11070864
8. Шелухин О.И., Барков В.В., Симонян А.Г. Обнаружение дрейфа концепта при классификации мобильных приложений с использованием автокодировщиков // Наукоемкие технологии в космических исследованиях Земли. 2023. Т. 15. № 3. С. 20–29. DOI:10.36724/2409-5419-2023-15-3-20-29. EDN:KBWOOG
9. Шелухин О.И., Барков В.В., Маторин Ф.А. Повышение эффективности классификации противоправных и нежелательных приложений в условиях фонового трафика с помощью автокодировщиков // Вестник Санкт-Петербургского государственного университета технологии и дизайна: Серия 1. Естественные и технические науки. 2023. № 3. С. 159–165. DOI:10.46418/2079-8199_2023_3_25. EDN:RLBDBM
10. Ososkov G., Goncharov P. Shallow and deep learning for image classification // Optical Memory and Neural Networks. 2017. Vol. 26. Iss. 4. PP. 221–248. DOI:10.3103/S1060992X1704004X

11. Шелухин О.И., Зегжда Д.П., Раковский Д.И., Самарин Н.Н., Александрова Е.Б. Интеллектуальные технологии информационной безопасности. М.: Горячая линия – Телеком, 2023. 384 с.
12. Шелухин О.И., Ерохин С.Д., Барков В.В. Создание базы данных сетевого трафика для автоматизации классификации мобильных приложений под управлением операционной системы Android // Нейрокомпьютеры: разработка, применение. 2019. № 1. С. 40–51. DOI:10.18127/j19998554-201901-06. EDN:BDDXDT
13. Шелухин О.И., Барков В.В. Экспериментальные исследования и создание базы данных сетевого трафика мобильных устройств под управлением операционной системы Android // Фундаментальные проблемы радиоэлектронного приборостроения. 2018. Т. 18. № 4. С. 1011–1017. EDN:ZABZMT

References

1. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. The MIT Press, 2016. 800 p.
2. Hinton G.E., Osindero S., Teh Y.W. A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*. 2006;18(7): 1527–1554. DOI:10.1162/neco.2006.18.7.1527
3. Salakhutdinov R., Hinton G.E. Deep Boltzmann Machines. *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (Clearwater Beach, USA). Proceedings of Machine Learning Research, vol.5*. 2009. p.448–455.
4. Kuzmina M.G. Multilayered autoencoders in problems of hyperspectral image analysis and processing. *Preprint M.V. Keldysh IAM*. 2021;28:21. DOI:10.20948/prepr-2021-28
5. Kramer M.A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*. 1991;37(2) 233–243. DOI:10.1002/aic.690370209
6. Bengio Y., Lamblin P., Popovici D., Larochelle H. Greedy Layer-Wise Training of Deep Networks. *In: Advances in Neural Information Processing Systems (B. Schölkopf, J. Platt, T. Hoffman (eds.)*. Cambridge; 2007. p.153–160.
7. Windrim L., Ramakrishnan R., Melkumyan A., Murphy R.J., Chlingaryan A. Unsupervised feature-learning for hyperspectral data with autoencoders. *Remote Sensing*. 2019;11(7):864. DOI:10.3390/rs11070864
8. Sheluhin O.I. Barkov V.V. Simonyan A.G. Concept Drift Detection in Mobile Applications Classification Using Autoencoders. *H&ES Research*. 2023;15(3):20–29. (in Russ.) DOI:10.36724/2409-5419-2023-15-3-20-29. EDN:KBWOOG
9. Sheluhin O.I. Barkov V.V. Matorin F.A. Improving the classification of illegal and unwanted applications under background traffic conditions using autoencoders. *Bulletin of the St. Petersburg State University of Technology and Design: Series 1. Natural and technical sciences*. 2023;3:159–165 (in Russ.) DOI:10.46418/2079-8199_2023_3_25. EDN:RLBDBM
10. Ososkov G., Goncharov P. Shallow and deep learning for image classification. *Optical Memory and Neural Networks*. 2017;26(4):221–248. DOI:10.3103/S1060992X1704004X
11. Sheluhin O.I., Zegzhda D.P., Rakovsk, D.I., Samari, N.N., Aleksandrova E.B. *Intelligent Technologies of Information Security*. Moscow: Goryachaya Liniya – Telecom Publ.; 2023. 384 p. (in Russ.)
12. Sheluhin O.I., Erokhin S.D., Barkov V.V. Creation of a Network Traffic Database for Automating the Classification of Mobile Applications under the Android Operating System. *Neurocomputers: Development, Application*. 2019;1:40–51. (in Russ.) DOI:10.18127/j19998554-201901-06. EDN:BDDXDT
13. Sheluhin O.I., Barkov V.V. Experimental Studies and Creation of a Network Traffic Database of Mobile Devices under the Android Operating System. *Fundamental Problems of Radio Electronic Instrument Engineering*ю 2018;18(4):1011–1017. (in Russ.) EDN:ZABZMT

Статья поступила в редакцию 30.10.2024; одобрена после рецензирования 25.11.2024; принята к публикации 12.12.2024.
The article was submitted 30.10.2024; approved after reviewing 25.11.2024; accepted for publication 12.12.2024.

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

МАТОРИН Фёдор Андреевич — аспирант кафедры «Информационная безопасность» Московского технического университета связи и информатики
https://orcid.org/0009-0002-4897-2338

Авторы сообщают об отсутствии конфликтов интересов.
The authors declare no conflicts of interests

ДЛЯ ЗАМЕТОК

СПбГУТ)))

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича

НТК ППС

65-я научно-техническая конференция профессорско-преподавательского состава, научных работников и аспирантов

ОСНОВНЫЕ НАУЧНЫЕ НАПРАВЛЕНИЯ:

- Инфокоммуникационные сети и системы
- Информационные технологии и программная инженерия
- Социальные технологии и экономика данных
- Радиоэлектронные системы и робототехника
- Кибербезопасность
- Сети связи специального назначения

17–21 ФЕВРАЛЯ 2025

apino.sut.ru

22 31 десятилетие науки и технологий

22—25 апреля 2025

СВЯЗЬ

37-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

Под патронатом
ТПП РФ

В рамках
Российская неделя высоких технологий РНВТ

План издания научной литературы 2024 г., п. 12

Усл.-печ. л.	Формат	Заказ
15,5	$60\times84_{1/8}$	№ 1616

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