ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

4(119)/2022

doi:10.31799/1684-8853-2022-4 ISSN 1684-8853 (print); ISSN 2541-8610 (online)

SCIENTIFIC JOURNAL

4(119)/2022

PEER REVIEWED JOURNAL

INFORMATSIONNO-UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder
A. Vostrikov
Publisher
Saint Petersburg State University of Aerospace Instrumentation
Editor-in-Chief
E. Krouk
Dr. Sc., Professor, Moscow, Russia
Executive secretary
O. Muravtsova
Editorial Board
S. Andreev
Dr. Sc., Tampere, Finland
V. Anisimov
Dr. Sc., Professor, Saint Petersburg, Russia
B. Bezruchko
Dr. Sc., Professor, Saratov, Russia
N. Blaunstein
Dr. Sc., Professor, Beer-Sheva, Israel
M. Buzdalov,
PhD, Researcher, Saint Petersburg, Russia
C. Christodoulou
PhD, Professor, Albuquerque, New Mexico, USA
A. Dudin
Dr. Sc., Professor, Minsk, Belarus
I. Dumer
PhD, Professor, Riverside, USA
M. Favorskaya
Dr. Sc., Professor, Krasnoyarsk, Russia
L. Fortuna
PhD, Professor, Catania, Italy
A. Fradkov
Dr. Sc., Professor, Saint Petersburg, Russia
A. Hramov
Dr. Sc., Professor, Innopolis, Russia
L. Jain
PhD, Professor, Canberra, Australia
G. Matvienko
Dr. Sc., Professor, Tomsk, Russia
A. Myllari
PhD, Professor, Grenada, West Indies
K. Samouylov
Dr. Sc., Professor, Moscow, Russia
J. Seberry
PhD, Professor, Wollongong, Australia
M. Sergeev
Dr. Sc., Professor, Saint Petersburg, Russia
A. Shalyto
Dr. Sc., Professor, Saint Petersburg, Russia
A. Shepeta
Dr. Sc., Professor, Saint Petersburg, Russia
Yu. Shokin
RAS Academician, Dr. Sc., Novosibirsk, Russia
A. Smirnov
Dr. Sc., Professor, Saint Petersburg, Russia
T. Sutikno
PhD, Associate Professor, Yogyakarta, Indonesia
Z. Yuldashev
Dr. Sc., Professor, Saint Petersburg, Russia
R. Yusupov
RAS Corr. Member, Dr. Sc., Professor, Saint Petersburg, Russia
A. Zeifman
Dr. Sc., Professor, Vologda, Russia
Editor: A. Larionova
Proofreader: T. Zvertanovskaia
Design: M. Chernenko, Yu. Umnitsyna
Layout and composition: Yu. Umnitsyna

Contact information
The Editorial and Publishing Center, SUAI
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia
Website: http://i-us.ru/en, e-mail: ius.spb@gmail.com
Tel.: +7 - 812 494 70 02

The Journal was registered in the Ministry of Press, Broadcasting and Mass Media of the Russian Federation.
Registration Certificate JD № 77-12412 from April, 19, 2002.
Re-registration in the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR) due to change of the founder A. Vostrikov,
PI № FS 77-82226 from November, 23, 2021
© A. Vostrikov, 2022

Submitted for publication 05.07.22. Passed for printing 09.09.22. Signed for printing: 13.09.22.
Format $60\times84_{1/8}$. Phototype CentSchbkCyrill BT. Digital printing.

Layout original is made at the Editorial and Publishing Center, SUAI.
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia
Printed from slides at the Editorial and Publishing Center, SUAI.
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia

The journal is indexed in Scopus.

Free distribution.

doi:10.31799/1684-8853-2022-4 ISSN 1684-8853 (print); ISSN 2541-8610 (online)
НАУЧНЫЙ ЖУРНАЛ

4(119)/2022

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель
А. А. Востриков
Издатель
Санкт-Петербургский государственный университет аэрокосмического приборостроения
Главный редактор
Е. А. Крук,
д-р техн. наук, проф., Москва, РФ
Ответственный секретарь
О. В. Муравцова
Редакционная коллегия:
С. Д. Андреев,
д-р техн. наук, Тампере, Финляндия
В. Г. Анисимов,
д-р техн. наук, проф., Санкт-Петербург, РФ
Б. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФ
Н. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, Израиль
М. В. Буздалов,
канд. техн. наук, научный сотрудник, Санкт-Петербург, РФ
Л. С. Джайн,
д-р наук, проф., Канберра, Австралия
А. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, Беларусь
И. И. Думер,
д-р наук, проф., Риверсайд, США
А. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФ
К. Кристодолу,
д-р наук, проф., Альбукерке, Нью-Мексико, США
Г. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФ
А. А. Мюлляри,
д-р наук, профессор, Гренада, Вест-Индия
К. Е. Самуйлов,
д-р техн. наук, проф., Москва, РФ
Д. Себерри,
д-р наук, проф., Волонгонг, Австралия
М. Б. Сергеев,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. В. Смирнов,
д-р техн. наук, проф., Санкт-Петербург, РФ
Т. Сутикноу,
д-р наук, доцент, Джокьякарта, Индонезия
М. Н. Фаворская,
д-р техн. наук, проф., Красноярск, РФ
Л. Фортуна,
д-р наук, проф., Катания, Италия
А. Л. Фрадков,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. Е. Храмов,
д-р физ.-мат. наук, Иннополис, РФ
А. А. Шалыто,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. П. Шепета,
д-р техн. наук, проф., Санкт-Петербург, РФ
Ю. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ
З. М. Юлдашев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Р. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
Редактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. Умницына

Адрес редакции: 190000, г. Санкт-Петербург,
ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: http://i-us.ru

Журнал входит в БД Scopus и 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.

Сдано в набор 05.07.22. Подписано в печать 09.09.22. Дата выхода в свет: 13.09.2022.
Формат 60×841/8. Гарнитура CentSchbkCyrill BT. Печать цифровая.
Усл. печ. л. 8,5. Уч.-изд. л. 11,7. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 436.
Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.
Отпечатано в редакционно-издательском центре ГУАП.
190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.
Распространяется бесплатно.

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-82226 от 23 ноября 2021 г.

© А. А. Востриков, 2022

УДК 621.391.01
doi:10.31799/1684-8853-2022-4-2-11

Научные статьи
Articles

Использование спектрального подхода при обработке изображений и произвольных данных

А. Л. Тимофеев[a], *канд. техн. наук, доцент, orcid.org / 0000-0003-2137-803X, a_l_t@inbox.ru*
А. Х. Султанов[a], *доктор техн. наук, профессор, orcid.org / 0000-0002-2830-3498*
И. К. Мешков[a], *канд. техн. наук, доцент, orcid.org /0000-0003-3479-3072*
А. Р. Гизатулин[a], *канд. техн. наук, доцент, orcid.org / 0000-0002-0753-0608*
[a]*Уфимский государственный авиационный технический университет, К. Маркса ул., 12, Уфа, 450008, РФ*

Введение: *рассматривается перенос операций обработки изображений в спектральную область. В силу дуальности представления информации кодирование, фильтрация, сжатие и другие преобразования возможны как в пространстве объекта, так и в области пространственных частот применительно к его пространственному спектру.* ***Цель:*** *оценить эффективность выполнения операций обработки изображений и произвольной цифровой информации в спектральной области.* ***Результаты:*** *предложен способ спектрального голографического кодирования, обеспечивающий выигрыш в помехоустойчивости на 7–8 дБ, обладающий меньшей вычислительной сложностью при изменении скорости кода в широких пределах. В процессе кодирования блок исходных данных заменяется кодовым словом, представляющим собой линейную одномерную голограмму виртуального точечного источника. Для передачи по каналу связи синтезируется сигнал с заданным спектром, форма которого является одномерной голограммой, значения которой округлены до одного бита, и представляет она собой последовательность нулей и единиц: единица означает наличие в спектре соответствующей гармоники, ноль — отсутствие. Для создания сигнала с таким линейчатым спектром достаточно сложить набор гармоник равной амплитуды с номерами, соответствующими номерам позиций единиц в голограмме. Эта операция является одним из видов мультиплексирования с ортогональным частотным разделением каналов, отличающимся тем, что частоты ортогональных поднесущих находятся в кратном отношении, а в качестве цифровой модуляции используется амплитудная манипуляция. Предложен спектральный метод сжатия изображений, основанный на детальном анализе пространственного спектра изображения и удалении из него большого количества малозначащих участков.* ***Практическая значимость:*** *сокращение объема информации, а соответственно, и размера изображения может составить 4–8 раз и более при незначительном снижении качества изображения.*

Ключевые слова *— голографическое кодирование, спектральная фильтрация, спектральное сжатие.*

Для цитирования: Тимофеев А. Л., Султанов А. Х., Мешков И. К., Гизатулин А. Р. Использование спектрального подхода при обработке изображений и произвольных данных. *Информационно-управляющие системы*, 2022, № 4, с. 2–11. doi:10.31799/1684-8853-2022-4-2-11

For citation: Timofeev A. L., Sultanov A. Kh., Meshkov I. K., Gizatulin A. R. Using the spectral approach in image and arbitrary data processing. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 2–11 (In Russian). doi:10.31799/1684-8853-2022-4-2-11

Введение

При обработке информации все шире применяются голографические методы, включающие как использование собственно голограмм, так и методы обработки цифровых массивов, не являющихся изображениями. При этом все виды цифровых голограмм образуют большие объемы данных, особенно если фиксируют большие углы обзора и объекты со значительной глубиной. И если голографические данные подлежат передаче или хранению, высокую важность приобретают средства кодирования и сжатия информации.

Существует большое число стандартов сжатия голографических изображений, в том числе JPEG, JPEG2000, Vp9 и HEVC/H.265 [1]. Во многих случаях методы сжатия изображений используют сходство между различными частями изображения во временной и частотной областях и сохраняют фрагменты похожих данных для уменьшения размера конечного файла с минимальным ущербом для качества. Кроме того, современные методы сжатия используют различные математические операции для коррекции ошибок. В работе [2] представлен гибридный алгоритм сжатия HEVC-Wavelet, который использует центральное ядро HEVC и канал 2D-вейвлета для прогнозирования ошибки сжатия.

В работе [3] рассмотрено современное состояние голографического кодирования данных и предложен вариант метода HEVC на основе направленного преобразования для повышения эффективности сжатия и кодирования. В [4] исследованы комбинации методов, состоящих из частотной фильтрации голограммы, разделения спектра Фурье отфильтрованной голограммы на действительно/мнимую и амплитудно-фазовую части, получения его вейвлет-разложения различными преобразованиями и дополнительной обработки вейвлет-коэффициентов.

Для эффективной обработки информации кроме задачи сжать голограммы необходимо обе-

спечить их надежную передачу по каналам связи и (или) хранение. Интересные возможности в этом плане предоставляет тот факт, что внутреннюю информационную избыточность голограмм можно использовать и для сжатия, и для повышения помехоустойчивости, причем в регулируемых соотношениях.

В статьях [5–7] рассмотрено использование голографического кодирования произвольных цифровых данных для повышения помехоустойчивости и надежности передачи информации. Дальнейшим расширением области применения голографических методов обработки информации является переход в спектральную область, спектральная голография. Непосредственная обработка пространственного спектра изображения представляет интерес для задач фильтрации, модуляции и кодирования. Одним из методов решения этих задач является введение дополнительной спектральной модуляции [8]. Поскольку речь идет об амплитудной и фазовой информации, решение таких задач для сигналов, изменяющихся во времени, обычно проводится с помощью голографических методов, в том числе с помощью спектральной голографии [9]. Спектральная модуляция применима, например, в беспроводных системах передачи информации, использующих сверхширокополосные шумоподобные сигналы [10].

В работе [11] предложено модулировать спектр цифровой информационной последовательности, представляющей сверхширокополосный шумоподобный сигнал, и восстанавливать информацию в приемнике с помощью спектральной обработки.

В статье [7] описано голографическое помехоустойчивое кодирование в пространстве виртуального объекта, носителя информации. Однако в силу дуальности представления информации аналогичные преобразования возможны и в частотной области, применительно к пространственному спектру объекта. Объединение двух методов — голографического кодирования и непосредственной обработки спектра — в метод спектрального голографического кодирования дает еще более высокий результат в повышении помехоустойчивости систем передачи и хранения как изображений, так и произвольной информации.

Спектральный подход позволяет повысить не только помехоустойчивость, но и эффективность сжатия изображений. Несмотря на наличие большого числа методов сжатия изображений [12–16], остается актуальной задача поиска специальных методов, обладающих большей совместимостью с методами помехоустойчивого кодирования. Переход из пространства изображения в область пространственных частот может дать возможность ее решить.

При рассмотрении сигналов, используемых как носители изображений, в алгоритмах сжатия с потерями для выбора оптимального набора параметров алгоритма для минимизации потерь необходимо учитывать характеристики изображений как в пространственной, так и в частотной области. Для этого можно использовать частотную декомпозицию (Фурье, вейвлет и др.) либо геометрическую интерпретацию [13]. Один из таких подходов — сжатие фазовых голограмм с использованием глубокого обучения нейросети — рассмотрен в работе [17].

Высокую эффективность с регулируемым уровнем вносимых искажений обеспечивает широко распространенный метод сжатия JPEG, использующий дискретное косинусное преобразование, реализуемое матрицей

$$DCT - 2_n = [\cos(k(l + 1/2)\pi/n)]_{0\le k,l<n}.$$

Для этого варианта дискретного косинусного преобразования с фиксированной размерностью вектора существуют алгоритмы, позволяющие свести количество операций умножения к минимуму.

Использование вейвлет-преобразований позволяет получить более высокую степень сжатия за счет удаления малозаметных деталей изображения. Большое значение имеет выбор метода пространственно-частотного разбиения спектра вейвлет-преобразования, например путем дальнейшей декомпозиции высокочастотных поддиапазонов для получения оптимального базиса [14]. Для повышения эффективности декомпозиции может быть использована адаптация базиса к содержанию изображения с количественной оценкой энтропии сигнала по базису вейвлет-пакетов y_k [15]:

$$H = -\sum_k y_k^2 \ln(y_k^2).$$

Кроме того, дополнительно повысить эффективность систем формирования и передачи цифровых изображений по каналам связи можно за счет комплексного использования алгоритмов сжатия и помехоустойчивого кодирования [18–20].

Таким образом, представляет интерес задача разработки методов помехоустойчивого кодирования и сжатия сигналов для повышения эффективности использования ресурсов канала связи (частотного ресурса, энергетики) при передаче голографических изображений путем частичного сокращения избыточности, содержащейся в голограмме, и использования оставшейся доли

избыточности для повышения помехоустойчивости. Далее рассматривается один из путей подхода к решению этой задачи — использование спектрального голографического кодирования и спектрального сжатия голограммы.

Спектральное кодирование произвольной цифровой информации

В описанном [7] методе голографического помехоустойчивого кодирования двоичный k-разрядный блок исходных данных заменяется N-разрядным ($N = 2^k$) кодовым словом, представляющим собой линейную одномерную голограмму виртуального точечного источника, позиция которого в виртуальном пространстве определяется значением кодируемого блока. В результате по цифровому каналу связи передается голограмма длиной N бит. В отличие от этого метода при спектральном голографическом кодировании используется передача по аналоговому каналу сигнала, форма спектра которого является той же цифровой одномерной голограммой, представляющей собой последовательность нулей и единиц — единица в i-й позиции означает наличие в спектре i-й гармоники, ноль — отсутствие.

Для создания сигнала $y(mT)$ с таким линейчатым спектром достаточно сложить набор гармоник равной амплитуды с номерами, соответствующими номерам позиций единиц в голограмме:

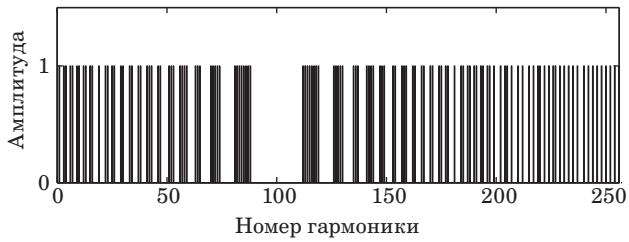
$$y(mT) = \sum_{i=1}^{N} (G(i) \times \times \sin(2\pi(m / M) \cdot i + r(i) \cdot 2\pi)),$$

где N — число гармоник; $G(i)$ — линейный массив голограммы; M — число отсчетов в сигнале; $r(i)$ — случайное число в диапазоне 0…1.

Гармоники с номерами, соответствующими позициям нулей, в образовании сигнала $y(mT)$ не участвуют. Эта операция является одним из видов мультиплексирования с ортогональным частотным разделением каналов (OFDM), отличающимся тем, что частоты N ортогональных поднесущих находятся в кратном отношении, а в качестве цифровой модуляции используется амплитудная манипуляция.

В результате синтезируется сигнал $y(mT)$ с линейчатым спектром $S(f)$, форма которого соответствует голограмме виртуального источника, порождаемого входным блоком данных (рис. 1). В данном примере число разрядов входного блока $k = 8$, кодированный сигнал $y(mT)$ содержит $N = 256$ гармоник, часть которых равна нулю.

Процедура синтеза сигнала $y(mT)$ аппаратно может быть реализована в аналоговой форме,



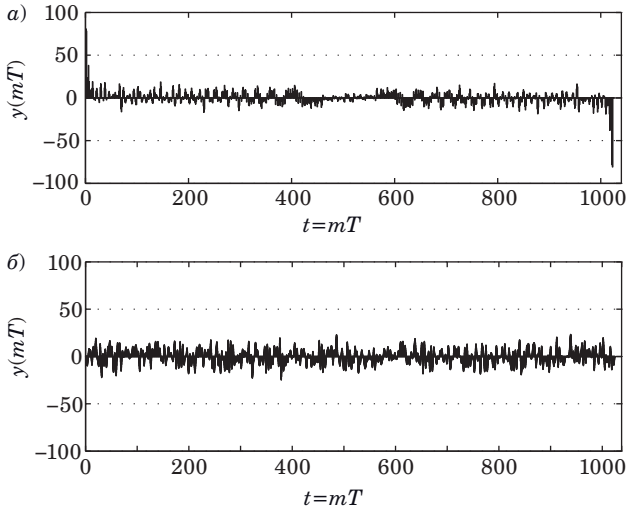
■ ***Рис. 1.*** Спектр-голограмма
■ ***Fig. 1.*** Spectrum-hologram

если будет жестко выдержано целочисленное отношение всех частот к первой гармонике. Однако во многих случаях точнее и проще проводить цифровой синтез, который может быть реализован двумя способами. Первый способ — алгебраическое сложение N гармоник, образующих спектр-голограмму, второй — обратное быстрое преобразование Фурье (БПФ) синтезированного спектра.

К длительности сигнала $y(mT)$ предъявляется одно требование — она должна быть не менее одного периода первой гармоники. Длительность большей величины не влияет на помехоустойчивость кодирования, но пропорционально увеличивает объем обрабатываемой информации.

Важной характеристикой является пик-фактор сигнала $y(mT)$, который достигает максимума при синфазности гармоник (рис. 2, *а*).

Повышенное значение пик-фактора предъявляет более высокие требования к линейности усилителя во избежание увеличения внеполосных излучений и снижения помехоустойчивости



■ ***Рис. 2.*** Форма сигнала: *а* — при синфазности гармоник; *б* — при случайной фазе гармоник
■ ***Fig. 2.*** Waveform: *а* — with common mode harmonics; *б* — at random phase of harmonics

канала связи [21–23]. Наилучший результат достигается при распределении фаз гармоник по случайному закону. Сигнал $y(mT)$ в этом случае имеет шумоподобную форму (рис. 2, *б*).

Таким образом, алгоритм спектрального кодирования заключается в следующем:

— по k-разрядному входному блоку данных формируется цифровая одномерная N-разрядная голограмма по алгоритму, приведенному в [7];

— синтезируется сигнал $y(mT)$ с линейчатым спектром $S(f)$, форма которого соответствует полученной голограмме;

— для синтеза сигнала используются N ортогональных гармоник со случайными фазами;

— длительность сигнала устанавливается равной периоду нижней гармоники;

— синтезированный сигнал используется для передачи по каналу связи.

Структурная схема кодера приведена на рис. 3.

Для проведения обратного преобразования и восстановления исходного блока данных в приемнике сигнал оцифровывается и вычисляется его спектр. Выбор частоты дискретизации сигнала при наличии шума проводится с учетом не только ширины спектра, но и уровня шума в канале [24].

При этом для построения спектра используется нормированный по длительности фрагмент принятого сигнала, содержащий целое число периодов каждой гармоники: один период первой гармоники, два — второй и т. д. до N периодов гармоники с номером $N(f_N)$. Выполнение этого требования позволяет получить линейчатый спектр, не содержащий побочных гармоник. Данное требование ограничивает длительность сигнала $y(mT)$ снизу, но не устанавливает ограничения сверху.

Цифровой массив, представляющий спектр, рассматривается как одномерная голограмма исходного цифрового блока, и декодируется описанным в [7] голографическим методом — производится восстановление исходного блока данных по цифровой голограмме.

Алгоритм декодирования состоит в следующем:

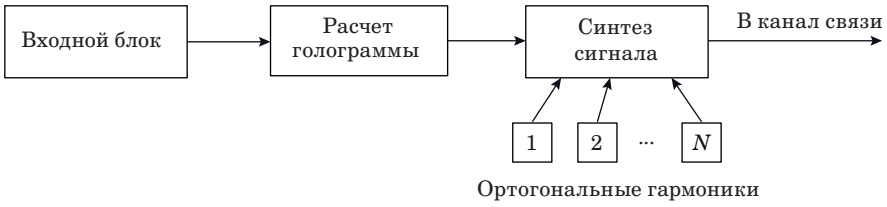
— производится аналого-цифровое преобразование (АЦП) с частотой дискретизации $f_д \geq 4f_N$;

— принятый сигнал нормируется по длине (выбирается фрагмент, строго равный длительности периода первой гармоники) и подвергается БПФ, в результате чего формируется голограмма передаваемого блока данных;

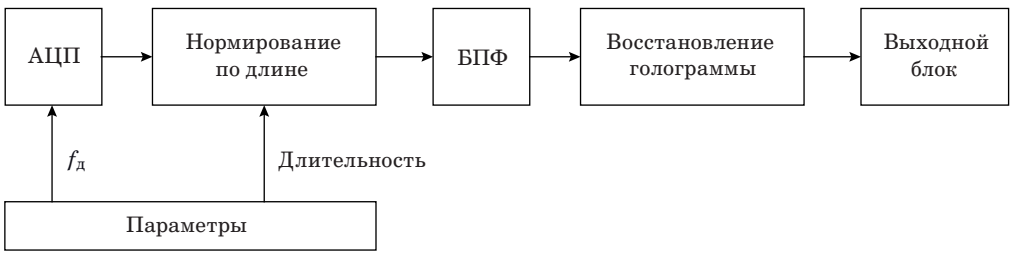
— проводится восстановление исходного блока по полученной голограмме по алгоритму, приведенному в [7].

Структурная схема спектрального декодера представлена на рис. 4.

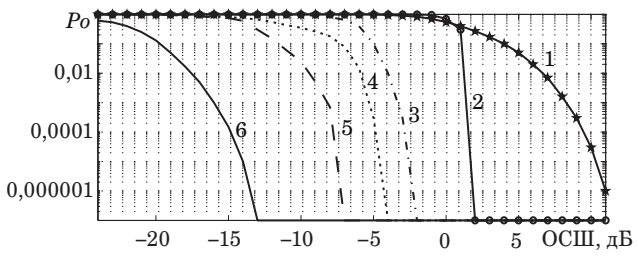
Перевод голографического кодирования из области времени в частотную область дает дополнительный выигрыш в помехоустойчивости. Моделирование в среде MATLAB процесса кодирования/декодирования в условиях наложения на сигнал аддитивного белого гауссова шума показало, что спектральный голографический код при размере блока данных $k = 8$ и использовании $N = 256$ гармоник в спектре сигнала обеспечивает вероятность ошибки декодирования 10^{-6} при отношении сигнал/шум (ОСШ) –13 дБ. Измеренные таким образом значения вероятности Po ошибки декодирования спектральным



■ ***Рис. 3.*** Структурная схема спектрального кодера
■ ***Fig. 3.*** Structural diagram of a spectral coder



■ ***Рис. 4.*** Структурная схема спектрального декодера
■ ***Fig. 4.*** Structural diagram of a spectral decoder



■ ***Рис. 5.*** Зависимость вероятности ошибки декодирования *Po* от ОСШ: *1* — без кодирования; *2* — РС-код; *3* — РМ-код; *4* — мажоритарный код; *5* — голографический код; *6* — спектральный код

■ ***Fig. 5.*** Dependence of decoding error probability *Po* on signal-to-noise ratio: *1* — no coding; *2* — PC code; *3* — PM code; *4* — majority code; *5* — holographic code; *6* — spectral code

кодом сопоставлены с полученными [7] зависимостями вероятности ошибки декодирования от ОСШ для кода Рида — Соломона (РС-кода), кода Рида — Маллера (РМ-кода), мажоритарного кода и голографического кода. Моделирование проводилось для 8-разрядного слова исходных данных при длине кодового слова 256 бит (скорость всех кодов $R = 1/32$) (рис. 5).

Из графиков видно, что спектральное голографическое кодирование обеспечивает выигрыш в помехоустойчивости 7–8 дБ по сравнению с кодированием во времени.

Другим преимуществом спектрального кода является меньшая вычислительная сложность для широкого диапазона скоростей кода. Декодирование широко применяемого кода Рида — Соломона представляет собой довольно сложную задачу, для решения которой разработано несколько видов алгоритмов. Например, алгоритм Питерсона — Горенстейна — Цирлера сводит задачу нахождения позиций и значений t ошибок к решению двух систем линейных уравнений порядка t. Для решения можно воспользоваться методом Гаусса, и тогда сложность вычислений будет иметь порядок t^3 [7]. Декодирование голографического кода в области времени состоит в N^2-кратном вычислении целочисленных сумм, что существенно проще алгоритмически и требует меньших вычислительных ресурсов. Для декодирования спектрального кода к этим операциям добавляется выполняемое один раз БПФ.

Спектральная обработка изображений

Операции с пространственным спектром изображений позволяют проводить фильтрацию и коррекцию изображений (размытие, повышение резкости, регулировку яркости и контрастности и т. д.), а также алгоритмически более сложные операции, такие как сжатие изображений.

Наиболее просто реализуется фильтрация узкополосных помех — необходимо провести дискретное преобразование Фурье, в полученном цифровом спектре обнулить частоты, на которых присутствуют помехи, и выполнить обратное преобразование Фурье. Если задачу решать цифровыми фильтрами, при необходимости удалить несколько гармоник в разных частях спектра, суммарный порядок фильтров пропорционально возрастает, а вычислительные затраты для реализации спектральной фильтрации остаются теми же — прямое и обратное преобразование Фурье.

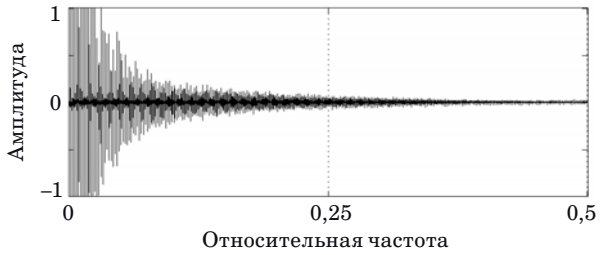
Спектральный метод сжатия изображений, основанный на анализе пространственного спектра изображения, отличается от существующих методов тем, что отсчеты спектра, подлежащие удалению, выбираются по структурным признакам — определенная часть регулярных интервалов спектра. Это позволяет при сжатии снизить объем вычислений практически до однократного вычисления спектра линейного массива.

Спектральное сжатие как голограмм, так и изображений удобнее оценивать на примере изображений. Для проведения спектральной обработки тестовое изображение размером 512×512 (рис. 6) путем построчной развертки преобразовано в линейный массив, спектр которого показан на рис. 7. Исследование изображения и его спектра, моделирование спектрального сжатия проведено в среде MATLAB.

Очевидный способ сокращения объема записанной информации, используемый практиче-

■ ***Рис. 6.*** Исходное изображение

■ ***Fig. 6.*** Original image

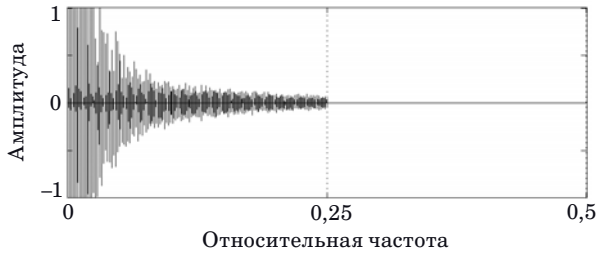


■ ***Рис. 7.*** Спектр линейного массива изображения
■ ***Fig. 7.*** Line array spectrum image

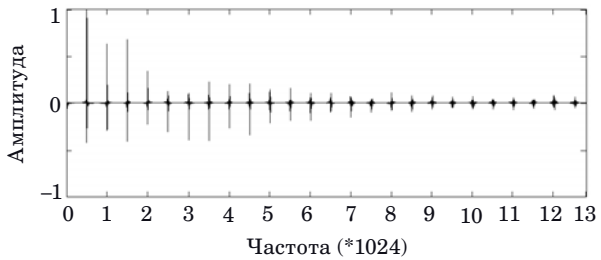
ски во всех кодеках, — стирание наименее значимой части спектра (в данном случае — высокочастотной), например, как показано на рис. 8, удаление высокочастотной половины спектра в диапазоне частот от $f_д/4$ до $f_д/2$.

Более подробное рассмотрение структуры спектра может дать дополнительные возможности по сокращению объема информации. При достаточно большом увеличении заметно, что наиболее значимые гармоники группируются по краям регулярных интервалов вокруг значений, кратных числу строк изображения (512) (рис. 9).

В каждом интервале из 512 частот центральная часть имеет низкий уровень, слабо влияет на качество полного изображения и поэтому может быть сокращена — на рис. 10 удалены две трети каждого интервала.



■ ***Рис. 8.*** Удаление высокочастотной части спектра
■ ***Fig. 8.*** Removal of the high-frequency part of the spectrum



■ ***Рис. 9.*** Увеличенная структура спектра
■ ***Fig. 9.*** Increased spectrum structure

Таким образом, удаление высокочастотной половины полного спектра и двух третей каждого интервала дает общее сокращение объема информации в изображении в 6 раз. При этом качество изображения практически не снижается (рис. 11).

Для формирования файла сжатого изображения необходимо после удаления фрагментов спектра уплотнить оставшиеся фрагменты и записать их единым массивом.

Таким образом, алгоритм спектрального сжатия включает в себя следующие операции:

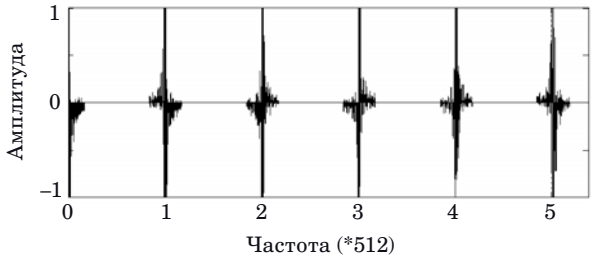
— выбор коэффициента сжатия;

— преобразование матрицы изображения в линейный массив;

— прямое БПФ;

— обнуление высокочастотной части спектра;

— обнуление средней части регулярных интервалов спектра. Размер удаляемой части определяется заданным коэффициентом сжатия;



■ ***Рис. 10.*** Удаления внутри регулярных интервалов спектра
■ ***Fig. 10.*** Deletes within regular intervals of the spectrum

■ ***Рис. 11.*** Сжатое в 6 раз изображение
■ ***Fig. 11.*** 6x compressed image

— удаление из массива обнуленных элементов, сокращение длины массива;

— запись в служебное поле массива параметров сжатия.

Алгоритм восстановления изображения:

— вставка нулевых значений массива по числу удаленных значений спектра;

— отражение спектра из диапазона относительных частот 0...0,5 в диапазон 0,5...1 (восстановление инверсной копии спектра);

— обратное БПФ;

— преобразование линейного массива в матрицу изображения.

Для проведения сравнительной оценки эффективности спектрального сжатия выполнено сжатие тестового изображения методом JPEG при коэффициенте сжатия 4,5. Искажение изображения, видимое при большом увеличении (рис. 12), оказалось больше, чем при спектральном сжатии с коэффициентом сжатия 6 (рис. 13).

■ ***Рис. 12.*** Сжатие JPEG в 4,5 раза (увеличенный фрагмент)

■ ***Fig. 12.*** JPEG compression 4.5 times (enlarged fragment)

■ ***Рис. 13.*** Спектральное сжатие в 6 раз (увеличенный фрагмент)

■ ***Fig. 13.*** Spectral compression by 6 times (enlarged fragment)

Для оценки искажений, вносимых при сжатии, вычислена с помощью встроенной функции пакета MATLAB среднеквадратическая ошибка $\varepsilon = \mathrm{mse}(\mathbf{a}, \mathbf{b})$, где **a** и **b** — матрицы исходного и сжатого изображений. Расчет показал, что среднеквадратическая ошибка при сжатии тестового изображения кодеком JPEG в 4,5 раза составила $\varepsilon_J = 16{,}4$, при спектральном сжатии в 6 раз среднеквадратическая ошибка $\varepsilon_S = 11{,}4$.

Заключение

При передаче изображений и произвольных цифровых данных по каналам связи большое значение имеет помехоустойчивое кодирование для защиты от широкополосного шума и узкополосных помех. Спектральный подход к кодированию и сжатию информации позволяет снизить требования к вычислительным ресурсам и использовать выигрыш от сокращения объема передаваемой информации для введения избыточности при помехоустойчивом кодировании и существенно повысить надежность передачи информации. Метод спектрального голографического кодирования, использующий цифровую голограмму для формирования спектра сигнала, в отличие от формирования сигнала, имеющего форму голограммы во временной области, обеспечивает выигрыш в помехоустойчивости 7–8 дБ. Метод спектрального сжатия, основанный на удалении регулярных фрагментов спектра, обеспечивает в ряде случаев сравнимую степень сжатия при внесении меньших искажений, чем кодек JPEG. Спектральный метод сжатия не является универсальным способом сжатия изображений. Он предназначен в первую очередь для сжатия голограмм, передаваемых по цифровым каналам связи. Специфика этой задачи заключается в том, что любая голограмма обладает высокой избыточностью, и для повышения скорости передачи информации необходимо сократить избыточность. Однако при сокращении избыточности уменьшается запас помехоустойчивости, поэтому алгоритм сжатия должен быть согласован с алгоритмом помехоустойчивого кодирования при ясном понимании того, что происходит со спектром сигнала. Поэтому и является эффективным совместное использование спектрального кодирования и спектрального сжатия.

Финансовая поддержка

Исследование выполнено за счет гранта Российского научного фонда № 22-29-00041, https://rscf.ru/project/22-29-00041/.

Литература

1. **Dufaux F., Xin Y., Pesquet-Popescu B., Schelkens P.** Compression of digital holographic data: An overview. *Proc. SPIE 9599*, *Applications of Digital Image Processing* XXXVIII, 95990I. doi:/10.1117/12.2190997
2. **Hajihashemi V., Najafabadi H. E., Gharahbagh A. A., Leung H., Yousefan M. Y., Tavares J. M.** A novel high-efficiency holography image compression method, based on HEVC, Wavelet, and nearest-neighbor interpolation. *Multimedia Tools and Applications*, 2021, no. 80, pp. 31953–31966. doi:10.1007/s11042-021-11232-0
3. **Peixeiro J. P., Brites C., Ascenso J. A., Pereira F.** Holographic data coding: benchmarking and extending HEVC with adapted transforms. *IEEE Transactions on Multimedia*, 2018, vol. 20, no. 2, pp. 282–297. doi:10.1109/TMM.2017.2742701
4. **Cheremkhin P. A., Kurbatova E. A.** Wavelet compression of off-axis digital holograms using real/imaginary and amplitude/phase parts. *Scientific Reports*, 2019, no. 9, pp. 7561. doi:/10.1038/s41598-019-44119-0
5. **Timofeev A. L., Sultanov A. Kh.** Holographic method of error-correcting coding. *Proc. SPIE 11146. Optical Technologies for Telecommunications* 2018, Proc., 2019, vol. 11146, 111461A. doi:10.1117/12.2526922
6. **Timofeev A. L., Sultanov A. Kh., Filatov P. E.** Holographic method for storage of digital information. *Proc. SPIE 11516. Optical Technologies for Telecommunications* 2019, 1151604, N.Y., SPIE, 2020. doi:10.1117/12.2566329
7. **Тимофеев А. Л., Султанов А. Х.** Построение помехоустойчивого кода на базе голографического представления произвольной цифровой информации. *Компьютерная оптика*, 2020, т. 44, № 6, с. 978–984. doi:10.18287/2412-6179-CO-739
8. **Зубов В. А.** Анализ изменяющихся во времени оптических сигналов и передаточных функций с использованием спектральной модуляции. *Квантовая электроника*, 1999, т. 29, № 2, с. 173–176.
9. **Mazurenko Yu. T.** Spectral holography. *Journal of Optical Technology*, 1994, vol. 61, no. 1, pp. 31–41.
10. **Narayanan R. M., Chuang J.** *Electron. Lett.*, 2007, vol. 43, no. 22, pp. 1211.
11. **Калинин В. И., Чапурский В. В.** Передача данных на основе шумовых сигналов со спектральной модуляцией. *Радиотехника и электроника*, 2015, т. 60, № 10, с. 1072–1082.
12. **Сэломон Д.** *Сжатие данных, изображений и звука*. М., Техносфера, 2006. 368 с.
13. **Гришенцев А. Ю.** Эффективное сжатие изображений на базе дифференциального анализа. *Журнал радиоэлектроники*, 2012, № 11. http://jre.cplire.ru/iso/nov12/1/text.html (дата обращения: 17.09.2018).
14. **Coifman R. R., Wickerhauser M. V.** Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory, Special Issue on Wavelet Transforms and Multires. Signal Anal,* 1992, no. 38, pp. 713–718.
15. **Умняшкин С. В., Гизатуллин Р. Р.** Сжатие изображений на основе блочной декомпозиции в области пакетного вейвлет-преобразования. *Цифровая обработка сигналов*, 2014, № 1, с. 46–51.
16. **Saupe D., Hamzaoui R., Hartenstein H.** Fractal image compression — An introductory overview. *Fractal Models for Image Synthesis, Compression, and Analysis*. D. Saupe, J. Hart (eds). ACM, New Orleans, SIGGRAPH'96 Course Notes 27. 66 p.
17. **Jiao S.** Compression of phase-only holograms with JPEG standard and deep learning. *Appl. Sci*, 2018, no. 8, p. 1258.
18. **Глумов Н. И.** Комплексный подход при выборе алгоритмов сжатия и помехоустойчивого кодирования для передачи цифровых изображений по каналам связи. *Компьютерная оптика*, 2004, № 26, с. 106–109.
19. **Гашников М. В., Глумов Н. И.** Повышение степени сжатия и визуального качества при иерархической компрессии изображений за счет предварительной фильтрации. *Компьютерная оптика*, 2005, № 28, с. 108–111.
20. **Демин В. В., Козлова А. В.** Методы кодирования-декодирования цифровых голограмм частиц. *Изв. вузов. Физика*, 2013, т. 5, № 10, с. 368–371.
21. **Fadeev D. K., Rashich A. V.** Optimal input power backoff of a nonlinear power amplifier for SEFDM-system. *Proc. of the NEW2AN, 2015 and 8th Conf.*, 2015, pp. 669–678.
22. **Tom A., Sahin A., Arslan H.** Suppressing alignment: Joint PAPR and out-of-band power leakage reduction for OFDM-based systems. *IEEE Transactions on Communications*, 2015, vol. 64, no. 3, pp. 1100–1109.
23. **Isam S., Darwazeh I.** Peak to average power ratio reduction in spectrally efficient FDM systems. *Proc. of the 18th Intern. Conf. on Telecommunications*, 2011. doi:10.1109/CTS.2011.5898951
24. **Тимофеев А. Л., Султанов А. Х.** Влияние шума и частоты дискретизации на погрешность дискретного представления изображений. *Информационно-управляющие системы*, 2021, № 5, с. 33–39. doi:10.31799/1684-8853-2021-5-33-39

UDC 621.391.01
doi:10.31799/1684-8853-2022-4-2-11

Using the spectral approach in image and arbitrary data processing

A. L. Timofeev[a], PhD, Tech., Associate Professor, orcid.org / 0000-0003-2137-803X, a_l_t@inbox.ru
A. Kh. Sultanov[a], Dr. Sc., Tech., Professor, orcid.org / 0000-0002-2830-3498
I. K. Meshkov[a], PhD, Tech., Associate Professor, orcid.org /0000-0003-3479-3072
A. R. Gizatulin[a], PhD, Tech., Associate Professor, orcid.org /0000-0002-0753-0608
[a]Ufa State Aviation Technical University, 12, K. Marks St., Ufa, 450008, Russian Federation

Introduction: The transfer of image processing operations to the spectral region is considered. Due to the duality of information representation, coding, filtering, compression and other transformations are possible both in the space of an object and in the region of spatial frequencies, in relation to its spatial spectrum. **Purpose:** To evaluate the efficiency of performing image processing operations and arbitrary digital information in the spectral domain. **Results:** A method for spectral holographic coding is proposed, which provides a gain in noise immunity by 7–8 dB and has less complexity of coding and decoding when changing redundancy over a wide range. During the coding process, the source data block is replaced by a code word, which is a linear one-dimensional hologram of a virtual point source. For transmission over a communication channel, we synthesize a signal with a given spectrum, the shape of which is a one-dimensional hologram, with its values rounded to one bit and the hologram itself being a sequence of zeros and ones – one means the presence of the corresponding harmonic in the spectrum, zero means the absence. To create a signal with such a line spectrum, it suffices to add a set of harmonics of equal amplitude with numbers corresponding to the position numbers of the units in the hologram. This operation is one of the types of orthogonal frequency division multiplexing, characterized in that the frequencies of the orthogonal subcarriers are in a multiple ratio, and amplitude shift keying is used as digital modulation. We propose a spectral method for image compression based on a detailed analysis of the spatial spectrum of an image and the removal of a large number of insignificant areas from it. **Practical relevance:** The reduction in the amount of information and, accordingly, the size of the image can be 4–8 times or more with a slight decrease in image quality.

Keywords — holographic coding, spectral filtering, spectral compression.

For citation: Timofeev A. L., Sultanov A. Kh., Meshkov I. K., Gizatulin A. R. Using the spectral approach in image and arbitrary data processing. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 2–11 (In Russian). doi:10.31799/1684-8853-2022-4-2-11

Financial support
This work was supported by the grant of Russian Science Foundation No. 22-29-00041.

References

1. Dufaux F., Xin Y., Pesquet-Popescu B., Schelkens P. Compression of digital holographic data: An overview. *Proc. SPIE 9599*, *Applications of Digital Image Processing* XXXVIII, 95990I. doi:/10.1117/12.2190997
2. Hajihashemi V., Najafabadi H. E., Gharahbagh A. A., Leung H., Yousefan M. Y., Tavares J. M. A novel high-efficiency holography image compression method, based on HEVC, Wavelet, and nearest-neighbor interpolation. *Multimedia Tools and Applications*, 2021, no. 80, p. 31953–31966. doi:10.1007/s11042-021-11232-0
3. Peixeiro J. P., Brites C., Ascenso J. A., Pereira F. Holographic data coding: benchmarking and extending HEVC with adapted transforms. *IEEE Transactions on Multimedia*, 2018, vol. 20, no. 2, pp. 282–297. doi:10.1109/TMM.2017.2742701
4. Cheremkhin P. A., Kurbatova E. A. Wavelet compression of off-axis digital holograms using real/imaginary and amplitude/phase parts. *Scientific Reports*, 2019, no. 9, pp. 7561. doi:/10.1038/s41598-019-44119-0
5. Timofeev A. L., Sultanov A. Kh. Holographic method of error-correcting coding. *Proc. SPIE 11146. Optical Technologies for Telecommunications* 2018, Proceedings Volume 11146, 111461A, 2019. doi:10.1117/12.2526922
6. Timofeev A. L., Sultanov A. Kh., Filatov P. E. Holographic method for storage of digital information. *Proc. SPIE 11516. Optical Technologies for Telecommunications* 2019, 1151604, N.Y., SPIE, 2020. doi:10.1117/12.2566329
7. Timofeev A. L., Sultanov A. Kh. Building a noise-tolerant code based on a holographic representation of arbitrary digital information. *Computer Optics*, 2020, vol. 44, no. 6, pp. 978–984 (In Russian). doi:10.18287/2412-6179-CO-739
8. Zubov V. A. Analysis of time-varying optical signals and transfer functions using spectral modulation. *Quantum Electronics*, 1999, vol. 29, no. 2, pp. 173–176 (In Russian).
9. Mazurenko Yu. T. Spectral holography. *Journal of Optical Technology*, 1994, vol. 61, no. 1, pp. 31–41.
10. Narayanan R. M., Chuang J. *Electron. Lett.*, 2007, vol. 43, no. 22, pp. 1211.
11. Kalinin V. I., Chapurski V. V. Data transmission on the basis of noise signals with spectral modulation. *Journal of Communications Technology and Electronics*, 2015, vol. 60, no. 10, pp. 1072–1082 (In Russian).
12. Salomon D. *Compression of data, images and sound*. Moscow, Tekhnosfera Publ., 2006. 368 p. (In Russian).
13. Grishentsev A. U. Efficient Image Compression Based on Differential Analysis. *Journal of Radio Electronics*, 2012, no. 11. Available at: http://jre.cplire.ru/iso/nov12/1/text.html (accessed 17 September 2018) (In Russian).
14. Coifman R. R., Wickerhauser M. V. Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory, Special Issue on Wavelet Transforms and Multires. Signal Anal*, 1992, no. 38, pp. 713–718.
15. Umniashkin S. V., Gisiatullin R. R. Compression of images based on block decomposition in the field of batch wavelet transform. *Digital Signal Processing*, 2014, no. 1, pp. 46–51 (In Russian).
16. Saupe D., Hamzaoui R., Hartenstein H. *Fractal image compression — An introductory overview*. In: *Fractal Models for Image Synthesis, Compression, and Analysis*. D. Saupe, J. Hart (eds). ACM, New Orleans, SIGGRAPH'96 Course Notes 27. 66 p.
17. Jiao S. Compression of phase-only holograms with JPEG standard and deep learning. *Appl. Sci*, 2018, no. 8, p. 1258.
18. Glumov N. I. An integrated approach to the selection of compression algorithms and error-correcting coding for the transmission of digital images via communication channels. *Computer Optics*, 2004, no. 26, pp. 106–109 (In Russian).
19. Gashnikov M. V., Glumov N. I. Improving the compression ratio and visual quality with hierarchical image compression due to pre-filtering. *Computer Optics*, 2005, no. 28, pp. 108–111 (In Russian).
20. Demin V. V., Kozlova A. V. Techniques of coding-decoding of particle digital holograms. *Russian Physics Journal*, 2013, vol. 56, no. 10, pp. 368–371 (In Russian).

21. Fadeev D. K., Rashich A. V. Optimal input power backoff of a nonlinear power amplifier for SEFDM-system. *Proc. of the NEW2AN, 2015 and 8th Conf.*, 2015, pp. 669–678.
22. Tom A., Sahin A., Arslan H. Suppressing alignment: Joint PAPR and out-of-band power leakage reduction for OFDM-based systems. *IEEE Transactions on Communications*, 2015, vol. 64, no. 3, pp. 1100–1109.
23. Isam S., Darwazeh I. Peak to average power ratio reduction in spectrally efficient FDM systems. *Proc. of the 18th Intern. Conf. on Telecommunications*, 2011. doi:10.1109/CTS.2011.5898951
24. Timofeev A. L., Sultanov A. Kh. Influence of noise and sampling rate on the discrete image representation error. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 33–39 (In Russian). doi:10.31799/1684-8853-2021-5-33-39

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

УДК 004.032
doi:10.31799/1684-8853-2022-4-12-19

Научные статьи
Articles

Оценка времени отклика среды для вычислений с интенсивным использованием данных

А. В. Горбунова[a], *канд. физ.-мат. наук, старший научный сотрудник, orcid.org/0000-0002-9183-0426, avgorbunova@list.ru*
В. М. Вишневский[a], *доктор техн. наук, профессор, orcid.org/0000-0001-7373-4847*
[a]*Институт проблем управления им. В. А. Трапезникова РАН, Профсоюзная ул., 65, Москва, 117997, РФ*

Введение: *объем цифровых данных непрерывно растет так же, как и потребность в их хранении и обработке в различных целях. Для проведения анализа данных используются высокопроизводительные вычислительные среды, связанные с методами распараллеливания, и, соответственно, приложения, интенсивно использующие данные. Отсутствие качественных инструментов оценки эффективности процесса параллельной обработки данных или задач приводит к избыточному выделению ресурсов.* ***Цель:*** *разработать математические модели сред для вычислений с интенсивным использованием данных и методы анализа их производительности, т. е. оценки среднего времени отклика системы на основе данных о производительности системы на уровне решения подзадач.* ***Результаты:*** *представлена математическая модель системы параллельных вычислений в виде системы массового обслуживания с параллельной обработкой заявок с различными вариантами архитектуры, в том числе с отличным от пуассоновского входящим потоком и неэкспоненциальным распределением времени обслуживания. В качестве метода анализа ее среднего времени отклика используется комбинация имитационного моделирования с одним из методов машинного обучения (искусственные нейронные сети). Эффективность метода подтверждается численными экспериментами и не зависит от типа входящего потока, типа распределения времени обслуживания заявок, а также от количества приборов в узлах системы. Погрешность аппроксимации среднего времени отклика не превышает 10 %, что позволяет оптимизировать общепринятую стратегию избыточного выделения ресурсов, значительно сократив их объем.* ***Практическая значимость:*** *представленные модели и метод их анализа могут быть использованы для эффективного планирования распределения ресурсов систем с интенсивным использованием данных.*

Ключевые слова *— приложения с интенсивным использованием данных, параллельные вычисления, система массового обслуживания, среднее время отклика, нейронные сети.*

Для цитирования: Горбунова А. В., Вишневский В. М. Оценка времени отклика среды для вычислений с интенсивным использованием данных. *Информационно-управляющие системы*, 2022, № 4, с. 12–19. doi:10.31799/1684-8853-2022-4-12–19
For citation: Gorbunova A. V., Vishnevsky V. M. Estimating the response time of a data-intensive computing environment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 12–19 (In Russian). doi:10.31799/1684-8853-2022-4-12–19

Введение

Согласно прогнозам одной из ведущих компаний в области цифровых технологий (IDC, International Data Corporation), к 2025 году глобальная сфера данных увеличится до 175 зеттабайт [1]. Это более чем в пять раз больше по сравнению с объемом цифровых данных, имевшихся по состоянию на 2018 год. При этом работа с большими данными стала доступной в основном благодаря развитию облачных технологий, а также появлению множества научных и технических приложений, пользующихся услугами облачных провайдеров для проведения вычислений и обработки данных [2–5].

Один из основных способов повышения производительности различного рода сервисов центров обработки данных заключается в распараллеливании вычислений [5]. К настоящему моменту разработано множество сред с параллельным подходом для ускорения работы приложений, интенсивно использующих данные [6].

В основе большинства сервисов центров обработки данных для проведения крупномасштабных вычислений находятся параллельные структуры. Они являются основным составным элементом процесса обработки данных систем параллельных и (или) распределенных вычислений. Однако в силу отсутствия хороших инструментов объективной оценки характеристик их производительности в настоящее время основной стратегией является избыточное выделение ресурсов, половина из которых простаивает большую часть времени [7]. Это приводит к значительному повышению затрат на содержание оборудования (серверов). Поэтому с точки зрения повышения производительности и эффективного использования ресурсов интерес представляет прогнозирование таких характеристик систем с интенсивным использованием данных, как ее среднее время отклика, в том числе в области высоких нагрузок.

В настоящей статье предлагается новый подход к оценке среднего времени отклика для высокопроизводительной вычислительной среды. Подход с использованием нейронных сетей (НС) позволяет довольно быстро и с приемлемой точностью получать оценки интересующих харак-

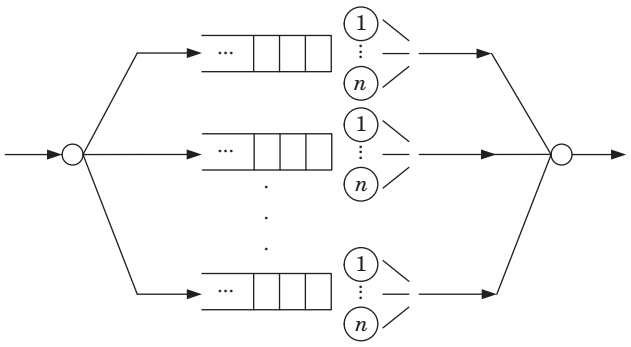
теристик. Преимущество подхода по сравнению с ранее известными заключается в универсальности, поскольку отсутствуют ограничения на архитектуру систем с параллельной обработкой заявок.

Система с параллельной обработкой заявок и известные методы ее анализа

В основе большинства центров обработки данных находятся параллельные структуры. Под параллельной структурой (системой) мы будем понимать систему массового обслуживания (СМО), каждый узел которой представляет собой самостоятельную СМО. При поступлении в систему задача (заявка) разбивается на части — независимые подзадачи (подзаявки), каждая из которых поступает на обслуживание на соответствующий узел (в подсистему). Задача считается выполненной после обработки последней из ее составляющих.

В классической СМО с параллельным обслуживанием заявок предполагается, что в каждой подсистеме имеется один обслуживающий прибор. Мы же расширим эту систему до самого общего случая, когда в каждой СМО может находиться по $n \geq 1$ приборов, т. е. каждая подсистема будет представлять собой СМО типа $G|G|n$ (рис. 1). Таким образом станет возможен анализ характеристик системы в условиях выделения дополнительных ресурсов. Результаты исследования позволят избежать избыточности в их выделении.

Одной из первых работ, посвященных анализу рассматриваемой СМО с подсистемами вида $M|M|1$, является статья [8]. Здесь было получено точное выражение для оценки среднего времени отклика системы в случае с двумя подсистемами $M|M|1$, $K = 2$. В большинстве следующих работ по данной тематике различными методами были получены только аппроксимации среднего времени отклика для $K > 2$. С подробным обзором публикаций можно ознакомиться, например, в [9, 10].



■ ***Рис. 1.*** Математическая модель системы с параллельной обработкой задач с подсистемами вида $G|G|n$

■ ***Fig. 1.*** Mathematical model of a system with parallel processing of tasks with subsystems of the form $G|G|n$

Среди основных приближенных методов исследований СМО с параллельным обслуживанием заявок с K подсистемами вида $M|M|1$ или $M|G|1$ можно перечислить матрично-геометрический подход; интерполяцию на основе данных, полученных в крайних случаях высоких и слабых входных нагрузок; подход с использованием элементов теории порядковых статистик; эмпирический подход (построение аналитических формул на основе данных, полученных с помощью симуляции) [8, 11–14].

Заметим, что точных решений для среднего времени отклика не существует даже в случае экспоненциального входящего и обслуживающих потоков при $K > 2$. Сложность анализа объясняется зависимостью времен пребывания подзаявок в подсистемах в силу их общих моментов поступления. Поэтому аналитический подход строится, как правило, на предположении об отсутствии этой зависимости. В результате применение полученных оценок ограничивается либо числом подсистем K, либо конкретным типом распределения обслуживающего потока, либо недостаточной точностью приближения. Это сужает область применения полученных аналитических выражений, в частности к современным системам, особенно если речь идет о распределениях с тяжелыми хвостами. Что касается случая с подсистемами более общего вида, т. е. $G|G|1$, то исследований на эту тему крайне мало, и появились они в основном в последнее время. Среди недавних работ, посвященных данной тематике, стоит отметить [15–19].

Система массового обслуживания с параллельным обслуживанием заявок является естественной моделью для многих реально существующих систем в различных областях, где осуществляется параллельная обработка заданий в целях повышения производительности. В данном случае речь идет не только о телекоммуникационных системах, но и о производственных сферах (сборка заказов, логистика и т. д.). Поэтому, несмотря на некоторое снижение активности исследований в этом направлении, их актуальность все еще велика [14].

Подход к исследованию СМО с параллельной обработкой заявок с использованием НС

Нейронные сети получили широкое распространение благодаря возможности их применения к решению слабо формализуемых задач.

Кроме того, они являются одним из основных инструментов анализа больших данных. Сложно в настоящее время назвать область, в которой НС или другие методы машинного обучения не нашли бы свое применение.

Математическая модель, которая используется для описания функционирования вычислительных сред с интенсивным использованием данных, представляет собой СМО. Методы машинного обучения и НС в частности к решению сложных задач теории массового обслуживания стали применяться относительно недавно, хотя данная перспектива была предсказуема. Под сложными задачами в области теории очередей понимаются задачи, решение которых невозможно получить с помощью известных аналитических методов, либо решение настолько сложно, что фактически получение численных результатов с помощью разработанных алгоритмов трудно реализуемо даже с учетом возможностей современных вычислительных машин. Обзор публикаций по применению методов машинного обучения к решению задач в области теории очередей можно найти в статье [20].

Идея применения НС базируется на возможности с их помощью решать задачу прогнозирования, т. е. задачу аппроксимации функции нескольких переменных. Понятно, что искомые оценки среднего времени отклика зависят, например, от значений таких параметров, как нагрузка системы, количество подсистем K и приборов в них, интенсивностей для входящего потока и времени обслуживания на приборах. Следовательно, задачу нахождения оценок времени отклика можно рассмотреть как задачу аппроксимации функции в зависимости от перечисленных параметров.

Для интерполяции функции необходим набор значений входных параметров и соответствующих им истинных значений аппроксимируемых величин. Получить истинные значения можно несколькими способами. Например, с помощью имитационного моделирования или с помощью точного аналитического решения, которое в данном случае отсутствует.

Необходимость в комбинации НС с имитационным моделированием или алгоритмическим решением возникает из-за значительных временных затрат, которые требуются при использовании только симуляции или только алгоритма. Поэтому мы ограничиваем количество входных данных, для которых необходимо использовать имитационное моделирование или вычислительный алгоритм. После чего на полученном наборе обучаем НС, которая позволит прогнозировать характеристики для любых промежуточных значений входных параметров без ограничений на их количество за минимальное время, сопоставимое со временем, необходимым для проведения расчетов по простой аналитической формуле.

В зависимости от доступной мощности вычислительной системы, программной среды для моделирования, уровня загрузки системы и т. д. время получения одного значения с помощью имитационного моделирования может варьироваться от нескольких десятков секунд до нескольких минут. При этом результат прогноза нейросети для заданного множества значений входных параметров выдается практически мгновенно. Таким образом, в зависимости от объема промежуточных значений, который ничем не ограничивается, время, затрачиваемое на оценивание искомых характеристик, значительно сокращается. Так, например, если количество промежуточных данных совпадает с количеством исходных, то общее время (с учетом имитационного моделирования данных для обучения нейросети) на оценку некоторой характеристики будет примерно в два раза меньше времени, потраченного на получение того же количества оценок, но только с помощью имитационной модели; если промежуточных данных в два раза больше исходных, то время сокращается в три раза и т. д.

Для обучения НС существует множество алгоритмов, большинство из которых реализовано в различных программных средах в виде готовых функций, как, например, в Python или MatLab. Поэтому наряду с написанием авторского программного кода одного из известных алгоритмов обучения можно воспользоваться готовым решением, что значительно ускоряет процесс, не требуя при этом слишком глубокого погружения в различные аспекты методов обучения нейросетей. При этом, как правило, в процессе любого обучения имеющаяся выборка исходных данных (имитационного моделирования) разбивается на тренировочную (и валидационную) выборку, на которой происходит непосредственное обучение нейросети одним из методов, и тестовую, на которой проверяется работоспособность обученной сети. Для более качественного анализа авторы проверяли работоспособность нейросети не только на тестовых данных, но и на существенном количестве промежуточных входных данных.

Среди основных преимуществ описанной методики можно выделить универсальность, так как имитационное моделирование иногда бывает единственно возможным способом анализа сложных систем. При этом оно является ресурсоемким инструментом. Благодаря применению нейросетей этот недостаток удается устранить.

Математическая модель параллельной вычислительной среды с интенсивным использованием данных

Более детально опишем архитектуру СМО с параллельным обслуживанием заявок, которую будем использовать для моделирования процесса функционирования параллельной вычислительной среды. Система состоит из K подсистем типа $G|G|n$ (см. рис. 1). Поступающая в систему задача, или в терминах теории очередей заявка, расщепляется на K подзаявок, каждая из которых встает в очередь в соответствующей подсистеме. Обслуживание в подсистемах происходит в порядке поступления подзаявок (дисциплина First In First Out, FIFO). В каждой подсистеме находится одинаковое число приборов n, $n \geq 1$. Выбор многолинейной системы позволит проанализировать повышение производительности узлов за счет выделения дополнительных ресурсов, поскольку таким образом моделируются дополнительные реплики серверов. Исходя из результатов исследования [17], ограничимся случаем с тремя репликами серверов, т. е. будем проверять точность предложенного подхода для $1 \leq n \leq 3$.

В качестве распределения времени обслуживания рассмотрим усеченное распределение Парето с плотностью распределения вида

$$f(x) = \frac{\alpha L^{\alpha} x^{-\alpha-1}}{1-(L/H)^{\alpha}},\ 0 \leq L \leq x \leq H,\ \alpha > 0.$$

Усеченное распределение Парето является трехпараметрическим. Параметр α — параметр формы, а параметры L и H фактически являются минимальным и максимальным значением случайной величины с данным распределением.

В статье [17] на основе эмпирических данных для поисковой системы Google предложены следующие значения параметров распределения времени обслуживания: $\alpha = 2{,}0119$, $L = 2{,}14$, $H = 276{,}6$, т. е. минимальное время обслуживания составляет 2,14 мс, а максимальное — 276,6 мс. Соответственно имеем, что среднее время обслуживания примерно равно 4,22 мс, дисперсия — 22,34, а коэффициент вариации CV, представляющий собой отношение корня из дисперсии к среднему значению случайной величины, будет составлять примерно 1,22.

Для распределения входящего потока рассмотрим несколько вариантов. Поскольку в отдельных исследованиях допускается предположение о пуассоновском характере входящего потока (CV = 1) для центров обработки данных, то не будем исключать экспоненциальное распределение для времени между соседними поступлениями заявок [17]. Также рассмотрим еще два типа распределения для входящего потока с коэффициентом вариации, отличным от единицы. В частности, для распределения Эрланга известно, что его коэффициент вариации всегда меньше единицы. Поэтому рассмотрим распределение Эрланга с плотностью вида

$$g(x) = \beta^2 x e^{-\beta x},\ x \geq 0, \beta > 0$$

и с CV $\approx$ 0,7. Также рассмотрим распределение с тяжелым хвостом (CV > 1), а именно гамма-распределение с плотностью

$$p(x) = \frac{\gamma^k}{\Gamma(\gamma)} x^{k-1} e^{-\gamma x},\ x \geq 0, \gamma > 0, k > 0,$$

для которого CV = 2.

Численный эксперимент

Проверим работоспособность и качество аппроксимации предложенного подхода с использованием НС. Первый вариант архитектуры высокопроизводительной вычислительной среды — это СМО с параллельной обработкой заявок с пуассоновским входящим потоком с интенсивностью λ, обратно пропорциональной среднему времени между соседними поступлениями заявок. Время обслуживания имеет усеченное распределение Парето с плотностью распределения $f(x)$ со средним значением $b = 4{,}22$ мс. Фактически каждая подсистема представляет собой СМО вида $M|G|n$.

Для обучения нейросети будем использовать входные данные, где коэффициент загрузки $\rho = \lambda b/n$ принимает значения на отрезке [0,1; 0,9] с шагом 0,1, число подсистем K меняется от двух до 24, а число приборов n — от одного до трех. В качестве выходных данных нейросети будет выступать среднее время отклика $E[R_K]$. Значения выходных параметров были получены с помощью имитационного моделирования, проведенного в программной среде Python.

Разобьем входные данные на два множества, соответствующие уровню слабой и высокой загрузок, т. е. для $\rho \in [0{,}1; 0{,}5]$ и $\rho \in [0{,}6; 0{,}9]$, и проведем обучение НС на каждом из этих двух наборов. При этом набор данных разбивается в соотношении 80 и 20 % на обучающую и тестовую выборки. По меркам НС нельзя сказать, что количество наборов данных для обучения велико. Тем не менее проверим точность работы предложенного подхода в данных условиях.

В качестве структуры нейросети выбран двухслойный персептрон с двумя скрытыми слоями по 10 нейронов в каждом с логистической функцией активации $\varphi(x) = 1/(1 + e^{-x})$. Обучение бу-

дем проводить методом обратного распространения ошибок или методом Адама в программной среде Python, где проводилось имитационное моделирование.

Для объективной проверки качества прогноза обученных НС будем использовать абсолютно незнакомые ей промежуточные данные. Для НС, обученной в области низких нагрузок, рассмотрим значения для прогноза $\rho \in [0{,}15;\ 0{,}55]$, а для НС, обученной в области более высоких нагрузок, — для $\rho \in [0{,}65;\ 0{,}85]$ с шагом 0,1 в обоих случаях.

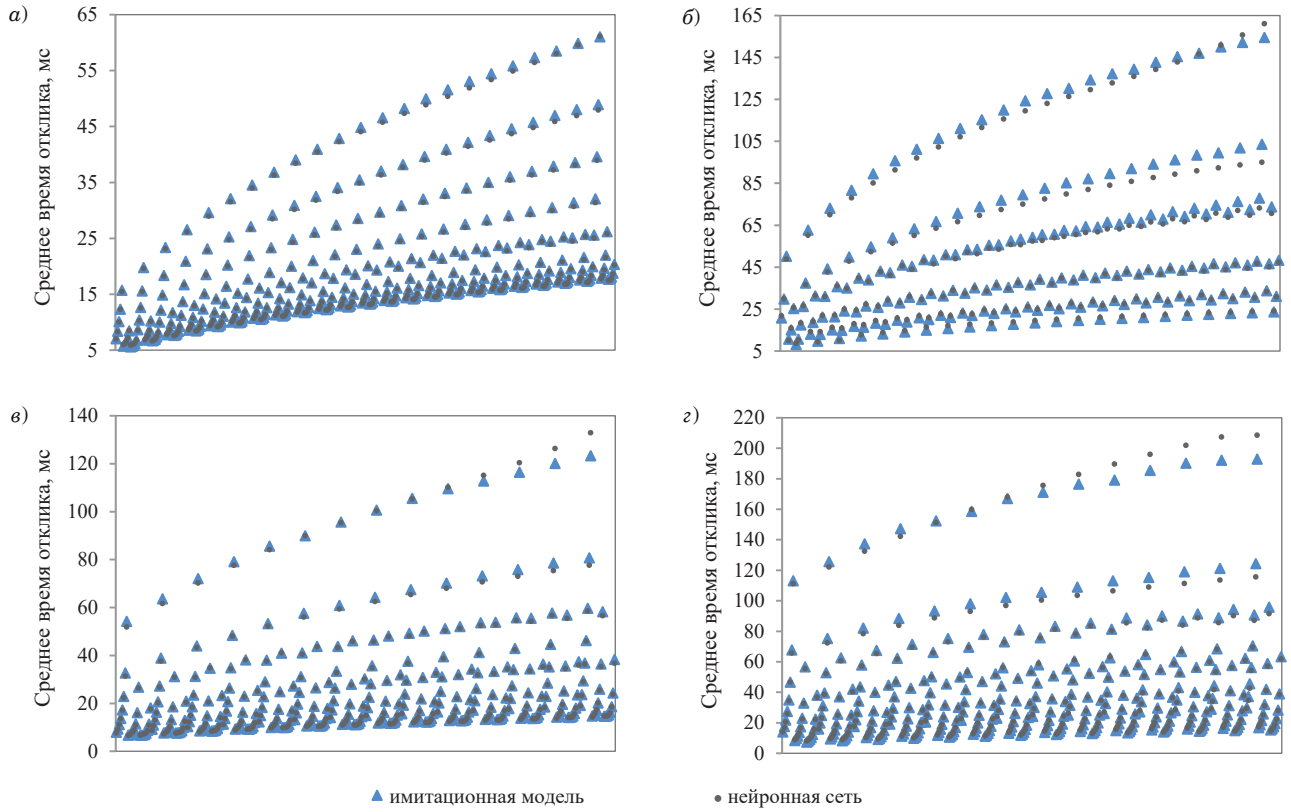
На рис. 2, *а–г* представлено отклонение оценок исследуемых характеристик от их истинных значений, полученных с помощью имитационного моделирования. Для большей детализации проанализируем относительную погрешность аппроксимации, а также ее среднее

$$\text{MAPE} = \frac{1}{n}\sum_{j=1}^{N}\left|\frac{y_j^* - y_j}{y_j}\right| \cdot 100\%,$$

максимальное и минимальное значения, где y_j^* — оценка исследуемой характеристики (математического ожидания времени отклика), полученная с помощью аналитических формул, а y_j — реальное значение оцениваемой характеристики, полученное в результате имитационного моделирования системы, $j = 1, \ldots, N$, N — количество наборов данных в выборке, предназначенной для оценки погрешности аппроксимации.

Примерно 99 % ошибок аппроксимации среднего времени отклика для $\rho \in [0{,}15;\ 0{,}55]$ не превышает 3 % (см. рис. 2, *а*). Для $\rho \in [0{,}65;\ 0{,}85]$ результат прогноза среднего времени отклика нейросетью оказался несколько хуже. В этом случае количество ошибок аппроксимации, не превышающее 7 %, составляет примерно 92 % от их общего количества (см. рис. 2, *б*). При этом максимальная ошибка аппроксимации не превышает 10 %.

Теперь проанализируем случай с распределением Эрланга для входящего потока с плотностью распределения $g(x)$. Для этого рассмотрим



■ ***Рис. 2.*** Сравнение результатов прогнозирования среднего времени отклика с результатами имитационного моделирования для СМО: *а* — с пуассоновским входящим потоком, $\rho \in [0{,}15;\ 0{,}55]$; *б* — с пуассоновским входящим потоком, $\rho \in [0{,}65;\ 0{,}85]$; *в* — с распределением Эрланга для входящего потока, $\rho \in [0{,}15;\ 0{,}85]$; *г* — с гамма-распределением для входящего потока, $\rho \in [0{,}15;\ 0{,}85]$

■ ***Fig. 2.*** Comparison of the results of predicting the average response time with simulation results for QS: *а* — with Poisson input, $\rho \in [0{,}15;\ 0{,}55]$; *б* — with Poisson input, $\rho \in [0{,}65;\ 0{,}85]$; *в* — with Erlang distribution for the incoming flow, $\rho \in [0{,}15;\ 0{,}85]$; *г* — with gamma distribution for the incoming flow, $\rho \in [0{,}15;\ 0{,}85]$

меньший набор входных данных для обучения, т. е. коэффициент загрузки принимает значения на отрезке [0,1; 0,9] с шагом 0,1, как и раньше, число приборов n от одного до трех, а число подсистем K меняется от двух до 16. Параметр распределения Эрланга определяется выражением $\beta = 2n\rho/b$.

После обучения нейросети строим соответствующие оценки для аналогичных промежуточных значений данных, т. е. для $\rho \in [0{,}15; 0{,}85]$ с шагом 0,1, $K = 2, \ldots, 16$ и $n = 1, 2, 3$. Так, для среднего времени отклика (см. рис. 2, *в*) примерно 98,5 % ошибок не превышает 5 %. Максимальная погрешность приближения $E[R_K]$ не превышает 8 %, что является приемлемым, особенно учитывая, что практически 99 % погрешностей находятся в пределах 5 %, о чем свидетельствуют их низкие средние значения (1,8 %).

Если интервалы между соседними поступлениями заявок имеют гамма-распределение $p(x)$ с параметрами $k = 0{,}25$ (CV = 2) и $\gamma = kn\rho/b$, то после обучения нейросети на тех же входных данных, что и в случае с распределением Эрланга, получим следующий результат для аналогичных значений промежуточных данных. Для среднего времени отклика (см. рис. 2, *г*) примерно 96 % относительных ошибок от их общего количества не превышает 6 %. При этом максимальная относительная погрешность приближения не превышает 10 %.

Если анализировать уровень снижения загрузки в случае выделения дополнительных серверов, то ситуация выглядит следующим образом. Вне зависимости от типа распределения для входящего потока из трех рассмотренных, при наличии в каждом узле двух серверов ($n = 2$) вместо одного происходит снижение среднего времени отклика примерно в два раза, что было ожидаемо. Если число серверов $n = 3$, то среднее время отклика уменьшается в среднем на 65 % по сравнению с $n = 1$. Расчет происходил для уровня высокой загрузки $\rho \in [0{,}6; 0{,}9]$, поскольку именно в этом случае выделение дополнительных ресурсов для повышения качества обслуживания является актуальным вопросом.

Резюмируя, можем сделать следующие выводы. Для 93 % входных данных погрешность аппроксимации среднего времени отклика не превышает 5 %. Поскольку мы рассматривали абсолютные значения относительной ошибки, то компенсация погрешности приближения в размере 5 % может в случае положительных оценок фактически привести к избыточному выделению ресурсов в размере 10 %. Однако по сравнению с принятым избыточным выделением ресурсов в размере 50 % от их общего объема для современных центров обработки данных экономия 40 % является серьезным преимуществом.

Обсуждение

В работе для оценки среднего времени отклика используются НС, однако предложенный подход не ограничивается только НС, можно применять и другие методы машинного обучения: градиентный бустинг, случайные деревья, бэггинг и др. При этом все сложности, связанные с обучением в данном случае НС, отражаются и на результатах прогнозирования. В частности, речь идет о выборе конкретного алгоритма обучения, подборе архитектуры НС, определении количества данных, необходимых для обучения, и пр. Тем не менее результаты численного эксперимента позволяют говорить о приемлемом качестве прогнозирования благодаря доступности широкого инструментария для обучения нейросетей в различных программных средах, что значительно снижает трудности и временные затраты. Кроме того, открываются возможности для оценки других характеристик рассматриваемой системы, например моментов времени отклика более высокого порядка.

Заключение

В работе предложен подход для оценки производительности систем с интенсивным использованием данных. С помощью комбинации имитационного моделирования с НС были получены оценки для среднего времени отклика. Преимущество подхода по сравнению с ранее известными заключается в универсальности, поскольку отсутствуют ограничения на архитектуру систем с параллельной обработкой заявок. Скорость работы обученной НС сопоставима с проведением вычислений по простой аналитической формуле в противовес имеющимся сложным вычислительным алгоритмам. При этом качество аппроксимации является довольно высоким и не зависит от архитектуры используемой математической модели.

Финансовая поддержка

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-29-06043.

Литература

1. **Reinsel D., Gantz J., Rydning J.** *IDC Report: The Digitization of the world from edge to core. IDC white paper*. Framingham, MA, International Data Corporation, 2018. 28 p.

2. **Han L., Ong H. Y.** Parallel data intensive applications using MapReduce: a data mining case study in biomedical sciences. *Cluster Computing*, 2015, vol. 18, pp. 403–418. doi:10.1007/s10586-014-0405-9
3. **Shanthi Thangam M., Vijayalakshmi M.** Data-intensive computation offloading using fog and cloud computing for mobile devices applications. *2018 Intern. Conf. on Smart Systems and Inventive Technology (ICSSIT)*, 2018, pp. 547–550. doi:10.1109/ICSSIT.2018.8748812
4. **Pandey A., Wang S., Calyam P.** Data-intensive workflow execution using distributed compute resources. *2019 IEEE 27th Intern. Conf. on Network Protocols (ICNP)*, 2019, pp. 1–2. doi:10.1109/ICNP.2019.8888119
5. **De Oliveira D. C., Liu J., Pacitti E.** *Data-intensive workflow management: For clouds and data-intensive and scalable computing environments*. Morgan & Claypool Publishers, 2019. 180 p. doi:10.2200/S00915ED1V01Y201904DTM060
6. **Khalid M., Yousaf M. M.** A comparative analysis of big data frameworks: An adoption perspective. *Applied Sciences*, 2021, vol. 11, no. 22, article number: 11033. doi.org/10.3390/app112211033
7. **Alesawi S., Nguyen M., Che H., Singhal A.** Tail latency prediction for datacenter applications in consolidated environments. 2*019 Intern. Conf. on Computing, Networking and Communications (ICNC)*, 2019, pp. 265–269. doi:10.1109/ICCNC.2019.8685505
8. **Nelson R., Tantawi A. N.** Approximate analysis of fork/join synchronization in parallel queues. *IEEE Transactions on Computers*, 1988, vol. 37, pp. 739–743. doi:10.1109/12.2213
9. **Thomasian A.** Analysis of fork/join and related queueing systems. *ACM Computing Surveys (CSUR)*, 2014, vol. 47, pp. 17:1–17:71. doi:10.1145/2628913
10. **Горбунова А. В., Зарядов И. С., Самуйлов К. Е., Сопин Э. С.** Обзор систем параллельной обработки заявок. *Discrete and Continuous Models and Applied Computational Science*, 2017, т. 25, № 4, с. 350–362. doi:10.22363/2312-9735-2017-25-4-350-362
11. **Varma S., Makowski A. M.** Interpolation approximations for symmetric fork-join queues. *Performance Evaluation*, 1994, vol. 20, pp. 245–265. doi:10.1016/0166-5316(94)90016-7
12. **Varki E., Merchant A., Chen H.** *The M/M/1 fork-join queue with variable subtasks*. https://www.cs.unh.edu/~varki/publication/2002-nov-open.pdf. (дата обращения: 24.03.2022).
13. **Zertal S., Harrison P.** Queueing models of RAID systems with maxima of waiting times. *Performance Evaluation*, 2007, vol. 64, pp. 664–689. doi:10.1016/j.peva.2006.11.002
14. **Sethuraman S.** *Analysis of fork-join systems: Network of queues with precedence constraints*. CRC Press, 2022. 104 p.
15. **Qiu Z., Perez J. F., Harrison P. G.** Beyond the mean in fork-join queues: Efficient approximation for response-time tails. *Performance Evaluation*, 2015, vol. 91, pp. 99–116. doi:10.1016/j.peva.2015.06.007
16. **Wang W., Harchol-Balter M., Jiang H., Scheller-Wolf A., Srikant R.** Delay asymptotics and bounds for multitask parallel jobs. *Queueing Systems*, 2019, vol. 91, pp. 207–239. doi:10.1007/s11134-018-09597-5
17. **Gorbunova A. V., Lebedev A. V.** Bivariate distributions of maximum remaining service times in fork-join infinite-server queues. *Problems of Information Transmission*, 2020, vol. 56, no. 1, pp. 73–90. doi:10.1134/S003294602001007X
18. **Nguyen M., Alesawi S., Li N., Che H., Jiang H.** A black-box fork-join latency prediction model for data-intensive applications. *IEEE Transactions on Parallel and Distributed Systems*, 2020, vol. 31, no. 9, pp. 1983–2000. doi:10.1109/TPDS.2020.2982137
19. **Gorbunova A. V., Lebedev A. V.** Response time estimate for a fork-join system with Pareto distributed service time as a model of a cloud computing system using neural networks. *Communications in Computer and Information Science*, 2022, vol. 1552, pp. 318–332. doi:10.1007/978-3-030-97110-6_25
20. **Вишневский В. М., Горбунова А. В.** Применение методов машинного обучения к решению задач теории массового обслуживания. *Информационные технологии и вычислительные системы*, 2021, № 4, с. 70–82. doi:10.14357/20718632210407

UDC 004.032
doi:10.31799/1684-8853-2022-4-12-19

Estimating the response time of a data-intensive computing environment

A. V. Gorbunova[a], PhD, Phys.-Math., Senior Researcher, orcid.org/0000-0002-9183-0426, avgorbunova@list.ru
V. M. Vishnevsky[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-7373-4847
[a]V. A. Trapeznikov Institute of Control Sciences of RAS, 65, Profsoyuznaya St., 117997, Moscow, Russian Federation

Introduction: The amount of digital data is constantly growing as well as the need for its storage and processing for various purposes. To conduct data analysis, high-performance computing environments associated with parallelization methods, and, accordingly, data-intensive applications are used. The lack of quality tools for evaluating the effectiveness of the process of parallel data processing or tasks leads to excessive allocation of resources. **Purpose:** To develop mathematical models of data-intensive computing environments and methods for their performance analysis, i.e., for estimating the average system response time based on the data on system performance at the level of subtask solving. **Results:** We present a mathematical model of a parallel computing system in the form of a queueing system

with parallel query processing on various architectures, including non-Poisson input flow and non-exponential service times. As a method for analyzing the average response time, we use a combination of simulation modeling with one of the machine learning methods (artificial neural networks). The effectiveness of the method is confirmed by numerical experiments and depends neither on the type of input flow, nor on the type of distribution of query service times, nor on the number of servers in the nodes of the system. The approximation error of the average response time does not exceed 10%, which makes it possible to optimize the generally accepted resource allocation, significantly reducing the amount of the resources. **Practical relevance:** The presented models and the method of their analysis can be used for efficient planning and allocation of resources for data-intensive systems.

Keywords — data-intensive applications, parallel computing, queueing system, average response time, neural networks.

For citation: Gorbunova A. V., Vishnevsky V. M. Estimating the response time of a data-intensive computing environment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 12–19 (In Russian). doi:10.31799/1684-8853-2022-4-12-19

Financial support

The reported study was funded by RFBR, project No. 19-29-06043.

References

1. Reinsel D., Gantz J., Rydning J. *IDC Report: The Digitization of the world from edge to core. IDC white paper*. Framingham, MA, International Data Corporation, 2018. 28 p.
2. Han L., Ong H. Y. Parallel data intensive applications using MapReduce: a data mining case study in biomedical sciences. *Cluster Computing*, 2015, vol. 18, pp. 403–418. doi:10.1007/s10586-014-0405-9
3. Shanthi Thangam M., Vijayalakshmi M. Data-intensive computation offloading using fog and cloud computing for mobile devices applications. *2018 Intern. Conf. on Smart Systems and Inventive Technology (ICSSIT)*, 2018, pp. 547–550. doi:10.1109/ICSSIT.2018.8748812
4. Pandey A., Wang S., Calyam P. Data-intensive workflow execution using distributed compute resources. *2019 IEEE 27th Intern. Conf. on Network Protocols (ICNP)*, 2019, pp. 1–2. doi:10.1109/ICNP.2019.8888119
5. De Oliveira D. C., Liu J., Pacitti E. *Data-intensive workflow management: For clouds and data-intensive and scalable computing environments*. Morgan & Claypool Publishers, 2019. 180 p. doi:10.2200/S00915ED1V01Y201904DTM060
6. Khalid M., Yousaf M. M. A comparative analysis of big data frameworks: An adoption perspective. *Applied Sciences*, 2021, vol. 11, no. 22, article number: 11033. doi.org/10.3390/app112211033
7. Alesawi S., Nguyen M., Che H., Singhal A. Tail latency prediction for datacenter applications in consolidated environments. 2*019 Intern. Conf. on Computing, Networking and Communications (ICNC)*, 2019, pp. 265–269. doi:10.1109/ICCNC.2019.8685505
8. Nelson R., Tantawi A. N. Approximate analysis of fork/join synchronization in parallel queues. *IEEE Transactions on Computers*, 1988, vol. 37, pp. 739–743. doi:10.1109/12.2213
9. Thomasian A. Analysis of fork/join and related queueing systems. *ACM Computing Surveys (CSUR)*, 2014, vol. 47, pp. 17:1–17:71. doi:10.1145/2628913
10. Gorbunova A. V., Zaryadov I. S., Samouylov K. E., Sopin E. S. A survey on queuing systems with parallel serving of customers. *Discrete and Continuous Models and Applied Computational Science*, 2017, vol. 25, no. 4, pp. 350–362 (In Russian). doi:10.22363/2312-9735-2017-25-4-350-362
11. Varma S., Makowski A. M. Interpolation approximations for symmetric fork-join queues. *Performance Evaluation*, 1994, vol. 20, pp. 245–265. doi:10.1016/0166-5316(94)90016-7
12. Varki E., Merchant A., Chen H. *The M/M/1 fork-join queue with variable subtasks*. Available at: https://www.cs.unh.edu/~varki/publication/2002-nov-open.pdf (accessed 24 March 2022).
13. Zertal S., Harrison P. Queueing models of RAID systems with maxima of waiting times. *Performance Evaluation*, 2007, vol. 64, pp. 664–689. doi:10.1016/j.peva.2006.11.002
14. Sethuraman S. *Analysis of fork-join systems: Network of queues with precedence constraints*. CRC Press, 2022. 104 p.
15. Qiu Z., Perez J. F., Harrison P. G. Beyond the mean in fork-join queues: Efficient approximation for response-time tails. *Performance Evaluation*, 2015, vol. 91, pp. 99–116. doi:10.1016/j.peva.2015.06.007
16. Wang W., Harchol-Balter M., Jiang H., Scheller-Wolf A., Srikant R. Delay asymptotics and bounds for multitask parallel jobs. *Queueing Systems*, 2019, vol. 91, pp. 207–239. doi:10.1007/s11134-018-09597-5
17. Gorbunova A. V., Lebedev A. V. Bivariate distributions of maximum remaining service times in fork-join infinite-server queues. *Problems of Information Transmission*, 2020, vol. 56, no. 1, pp. 73–90. doi:10.1134/S003294602001007X
18. Nguyen M., Alesawi S., Li N., Che H., Jiang H. A black-box fork-join latency prediction model for data-intensive applications. *IEEE Transactions on Parallel and Distributed Systems*, 2020, vol. 31, no. 9, pp. 1983–2000. doi:10.1109/TPDS.2020.2982137
19. Gorbunova A. V., Lebedev A. V. Response time estimate for a fork-join system with Pareto distributed service time as a model of a cloud computing system using neural networks. *Communications in Computer and Information Science*, 2022, vol. 1552, pp. 318–332. doi:10.1007/978-3-030-97110-6_25
20. Vishnevsky V. M., Gorbunova A. V. On the application of machine learning methods to solving problems queuing theory. *Journal of Information Technologies and Computing Systems*, 2021, no. 4, pp. 70–82 (In Russian). doi:10.14357/20718632210407

UDC 681.586.2
doi:10.31799/1684-8853-2022-4-20-28

Articles

Signal processing of capacitive force sensors installed in the foot of an anthropomorphic robot

K. D. Krestovnikov[a], *Junior Researcher, orcid.org/0000-0001-6303-0344*
A. A. Erashov[a], *Junior Researcher, orcid.org/0000-0001-8003-3643, erashov.a@iias.spb.su*
[a]*St. Petersburg Federal Research Center of the RAS, 39, 14th Line, 199178, Saint-Petersburg, Russian Federation*

Introduction: *The force-moment sensing of the functional surfaces in robots based on compact capacitive force sensors can significantly improve interaction with the environment and humans. Capacitive force sensors provide high measurement accuracy and speed, but the electromagnetic interference can affect significantly the signal. When processing signals the influence of external noise must be taken into account, which increases the computation time.* ***Purpose:*** *To apply the developed interface circuit for processing signals from capacitive primary force transducers in a real-world robot.* ***Results:*** *Experimental verification of the developed solutions implied simulation of a step of the pedipulator of an anthropomorphic robot with the calculation of the coordinates of the center of pressure exerted on the foot with the four installed capacitive sensors. Software filtering and measuring capabilities of the microcontroller made it possible to achieve a signal-to-noise ratio of approximately 62.24 dB, which allows a closed-loop control system to function correctly. The average time for calculating the coordinates of the center of pressure on the foot with software filtering of the signal on the on-board computer of the robot was from 3.1 to 6.1 ms and meets the requirements for the sensory system of a walking robot.* ***Practical relevance:*** *The interface circuit allows to scale the number of connected primary force transducers, while software processing allows to normalize transducer signals by applying the calculated correction factors. Proposed solutions can be used in different robotic systems for real-time force measurement.*

Keywords *— force sensors, capacitive sensors, humanoid robots, robotic sensing system, digital signal processing*

For citation: Krestovnikov K. D., Erashov A. A. Signal processing of capacitive force sensors installed in the foot of an anthropomorphic robot. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 20–28. doi:10.31799/1684-8853-2022-4-20-28

Introduction

The implementation of sensor systems of force-moment sensing into robotic systems makes it possible to improve their interaction with the environment and humans, which is currently an urgent scientific task. Anthropomorphic robots are one of the rapidly developing areas in robotics that require force-moment sensing. Implementation of gait and balance are actual basic tasks of anthropomorphic robotics. Stable gait of walking robots with constantly maintained balance on uneven surfaces requires information about the position of the zero moment point (ZMP) [1], as well as information about the distribution of pressure over the supporting surface of the robot foot. The robot's control system receives this information from the sensor system, which consists of sensors, interface circuits and computing devices. The quality of maintaining the balance and gait of the robot depends on the accuracy and processing speed of the sensor system. Its performance largely depends on the used communication interfaces, computing devices and software. The accuracy of the received data is influenced by the parameters of sensitivity, static and cyclic drift of the sensors, and the quality of the interface circuit.

Force-torque sensors for controlling the gait of robots can be placed in the intermediate joints of the robot foot [2, 3], as well as on the surface of the foot, which is in direct contact with the supporting surface [4, 5].

The first method of installing sensors is used in [2], where six-axis force sensors were used to control the gait of the anthropomorphic robot. For high-level operations on board the robot, an on-board computer is used. The functions of the lower-level control systems are performed by Infineon C167-CS microcontrollers. Controllers and on-board computer communicate via CAN bus. A similar method of installing force sensors was used in the feet of the anthropomorphic robot BHR-2 [3]. To process signals from the pedipulator sensors, a signal processor (DSP) is used, one of the ADC inputs of which receives a signal from a force sensor. CAN bus is used for communication between control devices. In further developments [6], the same method of installing force-torque sensors in the feet of the BHR-5 robot was chosen. Similar to the BHR-2 robot, the interface circuit of the sensor is installed in each foot of the pedipulators. RS422 is selected as the communication inter-face between the measuring circuit and the sensors, the PCI interface is used to exchange data with a PC running RTLinux. As the authors note, the modules of the developed system communicate in real time with a period of 2 ms.

In the pedipulators of the anthropomorphic LOLA robot, a cruciform six-axis force-torque sensor is installed between the foot and ankle [7]. Each beam of the cross contains one strain gauge

sensor. The lower-level controllers are responsible for data processing. They are represented by signal processors that transmit the processed data to the upper-level controller under the control of the QNX system. The controllers exchange data via a communication system Sercos-III based on Ethernet.

In the ankle joints of the pedipulators of robots KHR-2 [8] and KHR-3 [9] force-torque sensors are used to determine ZMP while the anthropomorphic robot is walking. The frequency of sending data from the sensors is 10 ms, the control cycle of the servo drives is 1 ms. An on-board computer running Windows XP with RTX (Real Time eXtension) generates control signals with a period of 10 ms.

The second method of placing the sensors is structurally simpler, since it does not require the installation of an intermediate link in the pedipulator and makes it possible to determine the distribution of contact forces along the foot of the robot. For walking robot sensing [4], the resistors located at the corners of the foot of a rectangular shape are used. The resistance of these sensors changes depending on the magnitude of the applied external force. On the foot of each pedipulator there are 4 such sensors, the signals of which are read by the ADC of the ATMEGA 128 microcontroller. The controller also calculates the ZMP of the foot, after which it transmits the data to the remote computer using the RS-485 interface. To filter the signals from the sensors, a third-order Chebyshev low-pass filter is used. Similarly, force sensors installed in the feet of the robot are used in [5, 10]. The signals from the sensors are converted using an ADC with a 10 ms sampling time. Measurements are performed in real time on a PC running the RTLinux operating system. As a force sensor for determining ZMP in [11] a piezoresistive strain gauge is used. As in the works mentioned above, the four sensors are installed on each foot of the robot at its corners.

To calculate the ZMP, a combination of several methods of measuring the forces applied to the robot pedipulator is possible. Thus in [12], the authors combine sensors installed on the surface of the foot and in the ankle of the robot. As well, in [13], the authors propose a combination of fluid pressure sensors in the drives and force sensors installed at the corners of the robot's foot.

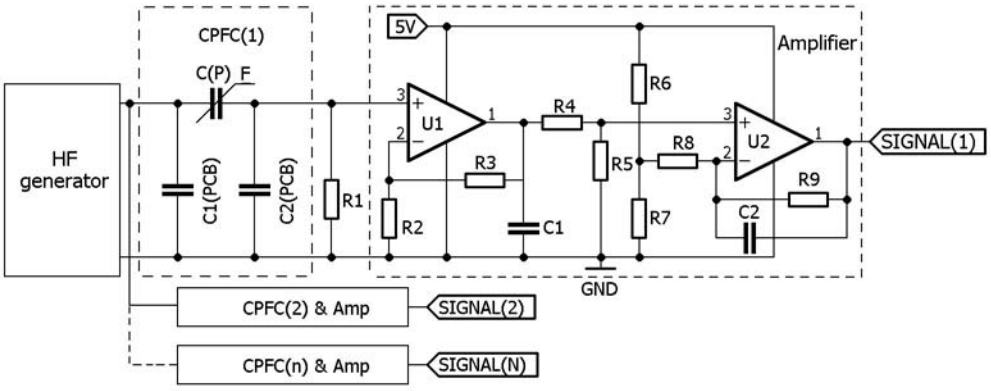
In contrast to the works discussed above, the matrix sensor [14] covers the entire foot surface of the H7 anthropomorphic robot pedipulators. The signal from the matrix sensor is measured with a period of 1 ms; an operational amplifier with an ADC is used as an interface circuit. The digitized data is transferred via the USB interface to a PC running RTLinux. A matrix of tensoresistive sensors with a dimension of 8×10 cells, covering the entire surface of the foot, is also used in [15].

Based on the analysis performed, as well as the data presented in the works [16–18], the time for obtaining data from the sensors and calculating the center of applied pressure area to the foot should be in the range of up to 10 ms for the robot's control system to respond sufficiently and to create a stable gait.

Installing force sensors directly on the foot surface, which contacts with the ground makes it possible to obtain information about the distribution of forces within the support area of the robot's feet and information about the properties of the supporting surface. Such primary force transducers are thin, so they do not affect robot's gait. This solution is easily scalable and can be effectively used in walking robots of medium and small size. The analysis of solutions described above showed that most of them are based on tensoresistive sensors with a film structure. The disadvantages of film strain gauges are high static and cyclic drift and a relatively short service life. In [19], we proposed a structure of sensor based on the capacitive principle, which do not have such disadvantages. In this work, we continue our research towards the use of capacitive force sensors in the foot of the anthropomorphic robot, with a focus on processing sensor signals. This work offers a comprehensive assessment of the practical application of the proposed sensors of our own design as part of the robotic system. The advantage of our work is that we do not consider individual tests of sensors with potential application in robotics, but test them on a real robot and process data from them for further use by the control system.

Circuit solutions of the sensors

In [19], we described in detail the structure of a capacitive primary transducer of pressure force and proposed some design ratios and an interface circuit for converting a signal into a dependence of voltage on the applied pressure force. The next stage of research was the development of a matrix of pressure force sensors based on the proposed sensor structure. The application of the interface scheme proposed in [19] for the matrix required a separate adjustment of each cell, which requires a significant amount of time. In [20], we proposed an interface circuit based on an operational amplifier, which requires tuning one matrix cell, and for the rest, the same component ratings are used. The disadvantage of the interface scheme described in [20] was that after adjusting the sensor the output signal from the interface circuit could change in a small range. This feature does not allow using the entire range of ADC values of the microcontroller, which negatively affects the accuracy of force measurement. In order to use the full range of ADC values, we have added



■ ***Fig. 1.*** Interface circuit of capacitive force sensors

a differential amplifier to our circuit design. A new schematic diagram for converting a capacitive primary force converter (CPFC) signal into a voltage is shown in Fig. 1.

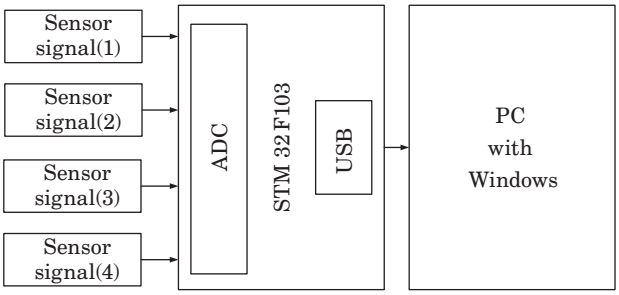
The principle of operation of the CPFC is based on a change in the distance between the electrodes of the cell, which changes in direct proportion to the applied force. When an external force is applied to the CPFC, its capacitance changes and, accordingly, the reactance changes as well. Since $C(P)$ and R_1 form a divider, changes in reactance $C(P)$ change the voltage applied to the noninverting input of the op-amp. A noninverting amplifier is built on U_1. Its gain is determined by the ratio of R_3 and R_2. The output signal from U_1 smoothed by C_1 is fed to the input of a differential amplifier built on U_2. The divider R_6R_7 forms the dead zone U_2, thereby excluding the DC component from the output signal of the sensor, which corresponds to the level of the output voltage from U_1 in the absence of an external force applied to the CPFC. Also, with the help of R_6R_7, it is possible to form a sensor dead zone at the beginning of the force measurement range. The amplification factor of the signal difference U_2 is determined based on the equation

$$k = R_5 / R_4, \quad (1)$$

where $R_5 = R_9$ and $R_4 = R_8$.

By changing the gain U_2, it is possible to increase the sensor's sensitivity relative to the value obtained after adjusting the gain U_1 and selecting R_1. The process of setting up the interface circuit without a differential amplifier and the calculated ratios are presented in detail in [20].

During experiments, four identical sensors were installed on the foot of the anthropomorphic robot. A significant interference in their output signals occurred, induced by servo drives and other robot electronics. To filter this noise, a capacitor C_2 was added to the feedback U_2, so the operational amplifier began to perform the function of a low-pass



■ ***Fig. 2.*** Block diagram of the connection of force sensors in the foot of the anthropomorphic robot

filter with a cutoff frequency of 15 Hz. The filter cutoff frequency was selected based on our experiments on gait simulation and the capabilities of the robot's mechanical subsystem. For robots in which the force applied to the foot can change at a frequency of more than 15 times per second, a higher cutoff frequency should be selected. The block diagram of connecting sensors to the robot's computing system is shown in Fig. 2.

The signals from the sensors are read out by means of a 12-bit microcontroller ADC, and then the received data is sent via the USB interface to the on-board computer of the robot for further processing.

Software processing of the sensors signals

Experiments with the developed sensor matrix of a similar design described in [20] showed that the levels of the output signals of the cells of the sensor matrix differ from each other. This feature is mainly associated with the imperfect manufacturing of the sensor itself. The deformable spacer may differ in thickness over the area of the matrix, which leads to differences in the capacities of the cells. The situation is similar with individual sensors installed in the foot of the robot. For correct subsequent calcu-

lations, normalization of sensors signals on primary signal processing level is required. The introduction of normalizing coefficients allows obtaining identical data from each of the sensors. We propose to use the following approach for this.

Sensors may differ from each other in sensitivity and in the initial level of the output signal. In the general case, the sensor output signal can be described by an equation of the straight form $y = kx + \Delta$. Thus, in order to bring the signals from each of the sensors to the same level, it is necessary to derive two correction factors k and Δ. To calculate Δ it is necessary to calculate the average level for the output signals E_{av}:

$$E_{av} = \sum_{i=1}^{n} E_i / n, \quad (2)$$

where n is the number of force sensor cells; E_i is the level of the output signal from each sensor in absence of external load.

Next, the correction value Δ_i is calculated for the output signal of each cell of the matrix:

$$\Delta_i = E_{av} - E_i. \quad (3)$$

Then, with a fixed equal force applied to all sensors, the proportional factor k_i is calculated for each sensor:

$$k_i = \frac{E_{av\_new} + \Delta_i}{E_{if}}, \quad (4)$$

where E_{av_new} is determined by (2) the average value of the output signal from the sensors at a fixed applied force and E_{if} is the output from the sensor cell when an external force is applied to it.

The normalized value of the signal from each sensor is determined by the following expression:

$$E_{i\_new} = k_i E_i + \Delta_i. \quad (5)$$

The introduction of these correction factors improves quality of the high-level software signal processing for the developed sensors.

To calculate the coordinates of the center of pressure (CoP) on the foot, one can use the following equations derived in [21]:

$$X_c = \frac{\sum_{i=1}^{N} X_i F_i}{\sum_{i=1}^{N} F_i}, \quad (6)$$

where X_i is the abscissa of the point of application of force to each support of the foot; F_i is the force applied to each support of the foot, and N it is the number of points to which the force is applied;

$$Y_c = \frac{\sum_{i=1}^{N} Y_i F_i}{\sum_{i=1}^{N} F_i}, \quad (7)$$

where Y_i is the ordinate of the point where the force is applied to each support of the foot. The proposed equations (6), (7) make it possible to calculate CoP for any number of sensors installed in arbitrary places on the robot's foot.

In accordance with the approach described above, the normalizing coefficients are calibrated according to equations (2)–(4), which are then stored in the memory of the on-board computer of the anthropomorphic robot. After that, it is possible to further receive data on the force from the microcontroller, normalize signals from each sensor according to equation (5) and process the obtained values. The result of applying the proposed approach is the coordinates of the center point of the applied pressure for each foot of the robot.

Experiments and results

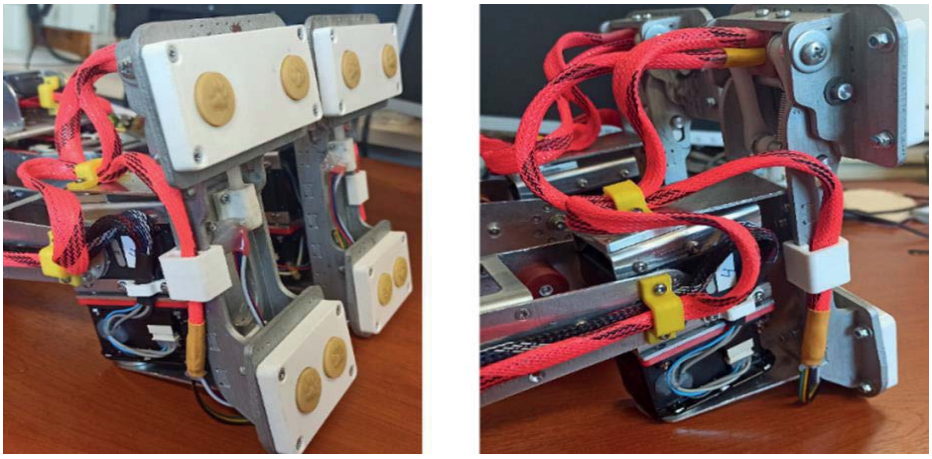
The proposed solutions were implemented on one pedipulator of a small-sized anthropomorphic robot with a total mass about 7 kg [22, 23]. The appearance of the prototype's foots are shown in Fig. 3.

The robot's foot has four reference points at each of which a sensor is installed, forming a trapezoid. Like a human foot, the forefoot of the robot can bend up to 39.5° and it is wider than the heel. Fig. 4 shows the appearance of the assembled printed circuit board, which implements the new interface circuit proposed in this work and two CPFC. These boards are installed on the toe and heel of the pedipulator foot.

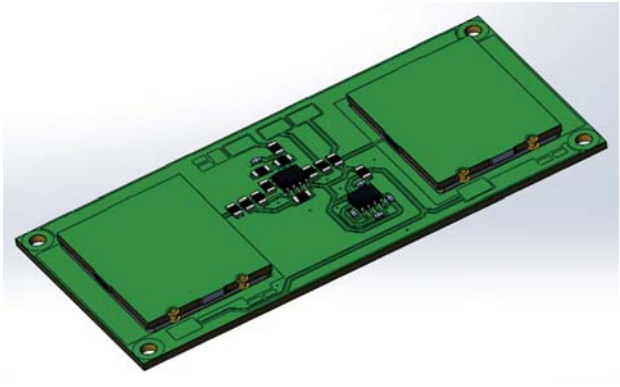
Software for processing data from foot sensors and visualization CoP were developed. Fig. 5 shows frames from the video [24], which demonstrates the movement of the pedipulator's foot from the state of full support on the surface to the state of support on the toe, as a result of which CoP also moves. In Fig. 5, CoP is shown with a red dot. Video [25] also demonstrates the CoP displacement to one of the four foot supports when force is applied to each sensor cell separately.

In the experiments, the median filter was used due to its computational simplicity and speed of operation [26], as well as the measurement capabilities of the ADC of the STM32F103C8T6 microcontroller.

The data for evaluating the noise were obtained as a result of measuring the signals from the sensors in the absence of external force and servo drives in the pedip-



■ ***Fig. 3.*** The feet of the anthropomorphic robot



■ ***Fig. 4.*** Sensors used in the foot of the anthropomorphic robot

ulator operating in mode of holding the moment. Three options for processing signals from sensors were considered: *case A* is measurements with a combination of increased ADC sampling time and median filter, *case B* is the measurements with increased sampling time (ADC parameter cycles is 7.5), *case C* is the measurements without filtering (cycles is 1.5). After the experiments, more than 100,000 values were obtained for each sensor in the foot. For each case under consideration, the standard deviations (STD) and average (AVG) values in the ADC values were calculated (Table).

It follows from Table that the filtering allows to reduce the level of noise in the output signals from the sensors. The use of a median filter with an increased sampling time of the ADC made it possible to reduce the noise level by an average of ≈ 5.33 times, while the maximum difference in the signals from the sensors was ≈ 33 %, and for the case without filtration ≈ 53 %.

The following series of experiments were carried out to evaluate the cumulative performance of the interface circuit computing system and the CPFC. For robotic control systems, an important parameter affecting the quality of work is time of receiving and processing data from the sensor system. The parameter of the total time for receiving and processing data is made up of the following time intervals:

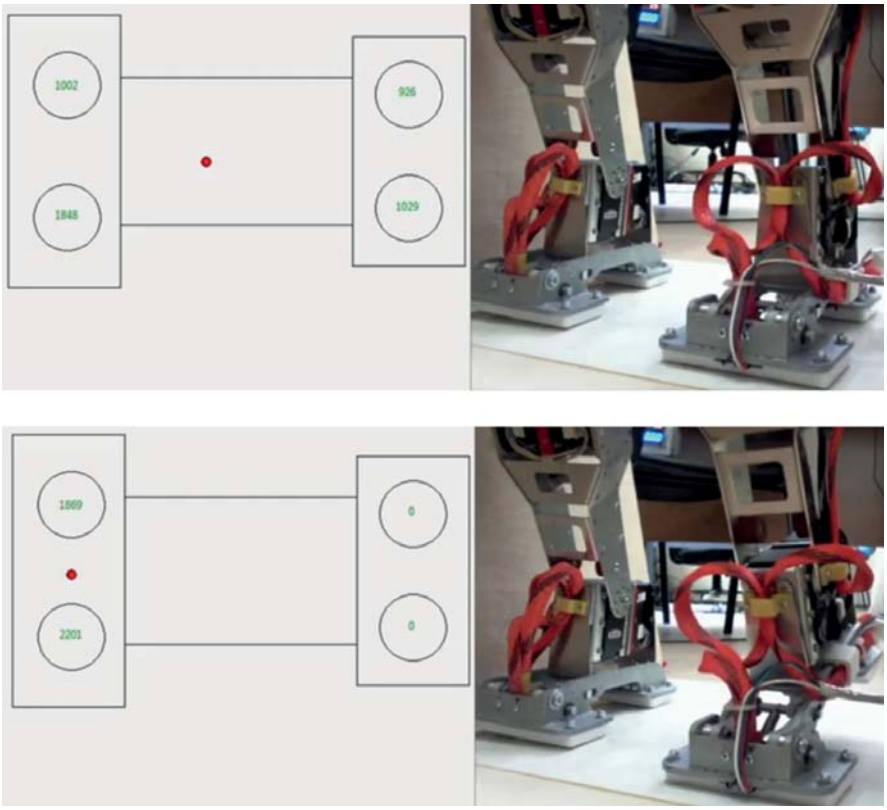
$$t_s = t_{reaction} + t_M + t_{transfer} + t_{processing}, \qquad (8)$$

where $t_{reaction}$ is the sensor response time and delay time from the operational amplifier; t_M is the cycle time of the microcontroller; $t_{transfer}$ is the time of data transfer over the communication interface, and $t_{processing}$ is time of data processing: normalization and calculation of the coordinates of the center of pressure.

Simulation of 10 steps of the pedipulator was performed with the force of foot pressure on the surface of 35 N. The process of simulation a step is shown in Fig. 6 [27].

In these experiments the total run time of each cycle of the microcontroller was measured using the built-in Data Watchpoint and Trace unit (DWT). The average cycle time of the microcontroller was t_M = 0.105 ms. It is worth noting that increasing the ADC sampling time does not significantly affect the overall cycle time of the microcontroller. According to the formula given in the documentation [28], measurement time increases from ≈ 1.167 μs to ≈ 1.429 μs, with cycles equal 1.5 and 7.5 respectively.

Time of data transfer from the microcontroller $t_{transfer}$ was obtained by measuring the time for performing the request-response operation on the on-board computer. The on-board computer sends a request to the microcontroller to receive a message with data on the force applied to each sensor. After receiving a response, the on-board computer records the time. Thus, $t_{transfer}$ is assumed to be equal to half of the time required for the request-response operation, and ranges approximately from 2 to 5 ms. Experiments with determining the point of the center of pressure of the foot showed that the use



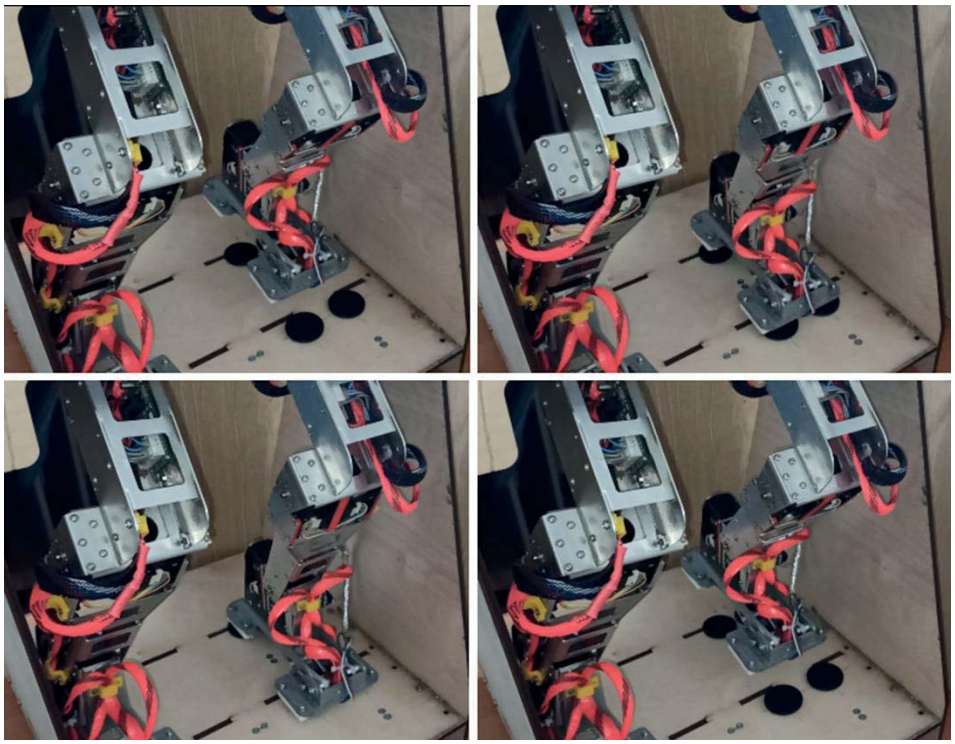
■ ***Fig. 5.*** Visualization of CoP at the different foot positions

■ Deviations of the sensor signals and the signal to noise ratio

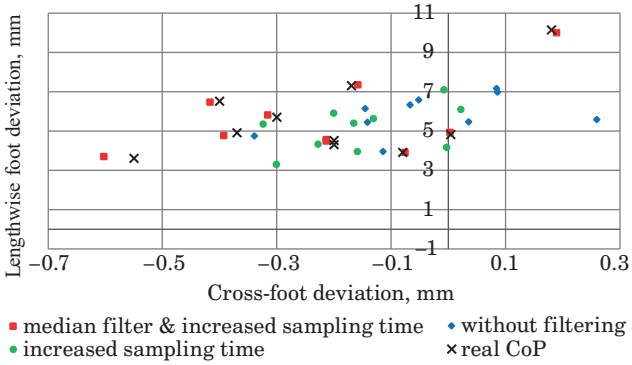
Parameter	Case	Sensor 1	Sensor 2	Sensor 3	Sensor 4	AVG
STD (ADC values)	A	2	2	3	2	2.25
	B	6	11	8	8	8.25
	C	11	19	9	9	12
Signal to noise ratio, dB	A	63.74	63.48	59.09	62.66	62.24
	B	57.41	48.49	52.21	53.43	52.89
	C	47.46	42.79	49.37	49.20	47.21

of the median filter had little effect on the signal processing time on the on-board computer. Using Python software, average time $t_{processing}$ for each case was ≈ 1 ms. Average time of calculating the coordinates of the center of pressure point t_s (8), considering the fact that the response time of each sensor is much less than the time of measurement by the microcontroller, as well as data transfer to the on-board computer and their processing, corresponds to the range from 3.1 to 6.1 ms. According to the results of the above analysis, the value of the CoP calculation time interval obtained using the sensors we proposed meets the requirements for the sensor system for the implementation of stable gait algorithms for anthropomorphic robots. The calculated values of the points of the center of pressure in the process of simulating the step of the pedipulator for different cases of signal filtering are shown in Fig. 7. The points are shown as deviations from the geometric center of the foot.

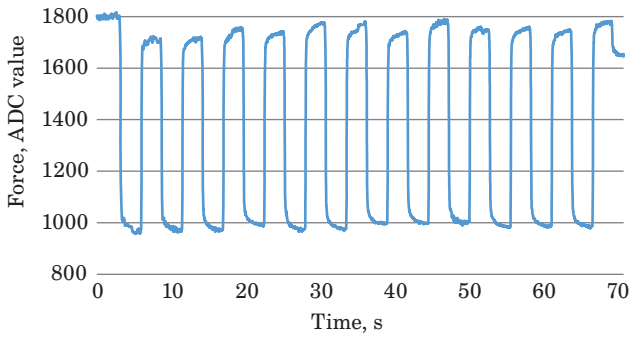
To determine the real CoP, 4 special platforms with tensoresistive sensors were installed in the experimental stand, on which the pedipulator stepped in the process of simulating a step (see Fig. 6). The measured forces in accordance with (6) and (7) were used to calculate the actual CoP value and are depicted by the symbol "×" in Fig. 7. As can be seen from the graph, the using of the median filter and the increased sampling time of the ADC allowed us to obtain CoP values closer to real ones. In this case,



■ ***Fig. 6.*** The one step of the pedipulator



■ ***Fig. 7.*** Average values of the center of pressure on the robot foot at 10 steps



■ ***Fig. 8.*** The change in force from one of the sensors when using a median filter and an increased ADC sampling time

on average, a deviation of ≈ 0.016 mm was obtained for the coordinates along the X-axis and 0.086 mm for the coordinates along the Y-axis.

The largest deviation between the points of the center of the applied pressure to the foot corresponds to ≈ 6.3 mm along the Y-axis, while the deviation was ≈ 4.5 % of the foot length. The scatter of points is affected by the inaccuracy of the positioning of the links and their non-rigidity, as well as backlash in the joints, which is confirmed by the graph of the simulation of steps in Fig. 8.

The graph in Fig. 8 corresponds to the signal from one of the sensors on the toe of the foot after using the median filter and the increased sampling time of the ADC. It can be seen, that the force applied to the sensor during the simulation of step was uneven, which confirms the above assumptions.

Conclusions

The paper presents the implementation of self-developed capacitive force sensors into the feet of a small-sized anthropomorphic robot. The offered force sensors are installed on the surface of the foot and allow calculating the center of applied pressure to the foot. The received data from the sensors can be used to calculate the ZMP [29]. An approach for primary

software signal processing has been developed, which allows reducing the difference in the characteristics of the sensors, occurring due to imperfections of CPFC manufacturing. The presented interface circuit with a differential amplifier allows one to use the entire ADC range and to increase the CPFC sensitivity.

During the experiments, the noise characteristics of the sensory system of the foot were considered when implementing various methods of filtering signals. The increase of sampling time of the ADC of the microcontroller in combination with the median filter allows obtaining an average signal-to-noise ratio of ≈ 62.24 dB without significant delays in signal processing.

The fundamental parameter for the correct operation of the system when the robot maintains balance or the implementation of a stable gait is the duration of the control cycle. This parameter is highly dependent on the speed of the sensor system. The average time for obtaining the coordinates of the center of pressure is in the range from 3.1 to 6.1 ms, which meets the requirements for cycle time of gait control systems for walking robots. The computation time can be reduced by using faster programming languages in the high-level part. In addition, real-time systems like QNX can achieve 2 ms USB transfer times without unpredictable delays.

The results obtained in the article show that the proposed solution can be applied in walking robots, including anthropomorphic ones for the implementation of stable gait algorithms. Besides the using of developed solutions can have significant influence on development of cyber-physical systems and collaborative robotics [30].

References

1. Kovalev A., Pavliuk N., Krestovnikov K., Saveliev A. Generation of walking patterns for biped robots based on dynamics of 3D linear inverted pendulum. *Intern. Conf. on Interactive Collaborative Robotics*, Aug. 2019, pp. 170–181. doi:10.1007/978-3-030-26118-4_17
2. Gienger M., Loffler K., Pfeiffer F. Towards the design of a biped jogging robot. *Proc. ICRA, IEEE Intern. Conf. on Robotics and Automation*, May 2001, vol. 4, pp. 4140–4145. doi:10.1109/ROBOT.2001.933265
3. Li J., Huang Q., Zhang W., Yu Z., Li K. Flexible foot design for a humanoid robot. *IEEE Intern. Conf. on Automation and Logistics*, 2008, pp. 1414–1419. doi:10.1109/ICAL.2008.4636375
4. Choi K. C., Lee H. J., Lee M. C. Fuzzy posture control for biped walking robot based on force sensor for ZMP. *SICE-ICASE Intern. Joint Conf.*, 2006, pp. 1185–1189. doi:10.1109/SICE.2006.315363
5. Erbatur K., Okazaki A., Obiya K., Takahashi T., Kawamura A. A study on the zero moment point measurement for biped walking robots. *IEEE 7th Intern. Workshop on Advanced Motion Control. Proc.*, July 2002, pp. 431–436. doi:10.1109/AMC.2002.1026959
6. Yu Z., Huang Q., Ma G., Chen X., Zhang W., Li J., Gao J. Design and development of the humanoid robot BHR-5. *Advances in Mechanical Engineering*, 2014, 6, p. 852937. https://doi.org/10.1155/2014/852937
7. Lohmeier S., Buschmann T., Ulbrich H. Humanoid robot LOLA. *IEEE Intern. Conf. on Robotics and Automation*, 2009, pp. 775–780. doi:10.1109/ROBOT.2009.5152578
8. Kim J. Y., Park I. W., Oh J. H. Experimental realization of dynamic walking of the biped humanoid robot KHR-2 using zero moment point feedback and inertial measurement. *Advanced Robotics*, 2006, vol. 20(6), pp. 707–736.
9. Park I. W., Kim J. Y., Oh J. H. Online walking pattern generation and its application to a biped humanoid robot – KHR-3 (HUBO). *Advanced Robotics*, 2008, vol. 22(2-3), pp. 159–190. doi:10.1163/156855308X292538
10. Prahlad V., Dip G., Meng-Hwee C. Disturbance rejection by online ZMP compensation. *Robotica*, 2008, vol. 26(1), p. 9. doi:10.1017/S0263574707003542
11. Yasin A., Xu Q., Chen B., Lu Q., Khan M. W. Design of a 23-DoF small humanoid robot with ZMP force sensors. *Intern. Conf. "Informatics in Control, Automation and Robotics"*, 2011, pp. 31–38. doi:10.1007/978-3-642-25899-2_5
12. Erbatur K., Seven U., Taskiran E., Koca O., Kiziltas G., Unel M., Onat A. SURALP-L-The leg module of a new humanoid robot platform. *Humanoids 2008 – 8th IEEE-RAS Intern. Conf. on Humanoid Robots*, pp. 168–173. doi:10.1109/ICHR.2008.4755963
13. Hyon S. H., Suewaka D., Torii Y., Oku N., Ishida H. Development of a fast torque-controlled hydraulic humanoid robot that can balance compliantly. *IEEE-RAS 15th Intern. Conf. on Humanoid Robots (Humanoids)*, 2015, pp. 576–581. doi:10.1109/HUMANOIDS.2015.7363420
14. Kagami S., Takahashi Y., Nishiwaki K., Mochimaru M., Mizoguchi H. High-speed matrix pressure sensor for humanoid robot by using thin force sensing resistance rubber sheet. *IEEE Sensors*, 2004, pp. 1534–1537. doi:10.1109/ICSENS.2004.1426481
15. Shimojo M., Araki T., Ming A., Ishikawa M. A ZMP sensor for a biped robot. *Proc. IEEE Intern. Conf. on Robotics and Automation*, 2006, pp. 1200–1205. doi:10.1109/ROBOT.2006.1641872
16. Yang U. J., Kim J. Y. Mechanical design of powered prosthetic leg and walking pattern generation based on motion capture data. *Advanced Robotics*, 2015, vol. 29(16), pp. 1061–1079. doi:10.1080/01691864.2015.1026939
17. Gorobtsov A. S., Andreev A. E., Markov A. E., Skorikov A. V., Tarasov P. S. Features of solving the inverse

dynamic method equations for the synthesis of stable walking robots controlled motion. *SPIIRAS Proc.*, 2019, vol. 18(1), pp. 85–122 (In Russian). doi:10.15622/sp.18.1.85-122

18. Kim I. S., Han Y. J., Hong Y. D. Stability control for dynamic walking of bipedal robot with real-time capture point trajectory optimization. *Journal of Intelligent & Robotic Systems*, 2019, vol. 96(3), pp. 345–361. doi:10.1007/s10846-018-0965-7

19. Krestovnikov K., Saveliev A., Cherskikh E. Development of a circuit design for a capacitive pressure sensor, applied in walking robot foot. *IEEE 20th Mediterranean Electrotechnical Conf. (MELECON)*, 2020, pp. 243–247. doi:10.1109/MELECON48756.2020.9140509

20. Krestovnikov K., Erashov A., Bykov A. Development of circuit solution and design of capacitive pressure sensor array for applied robotics. *Robotics and Technical Cybernetics*, 2020, vol. 8(4), pp. 296–307 (In Russian). doi:10.31776/RTCJ.8406

21. Kajita S., Hirukawa H., Harada K., Yokoi K. *Introduction to humanoid robotics*. Ser. Springer Tracts in Advanced Robotics, vol. 101. Springer, Berlin, Heidelberg, 2014. 222 p. doi:10.1007/978-3-642-54536-8

22. Pavluk N., Ivin A., Budkov V., Kodyakov A., Ronzhin A. Mechanical leg design of the anthropomorphic robot Antares. *Intern. Conf. on Interactive Collaborative Robotics*, 2016, pp. 113–123. doi:10.1007/978-3-319-43955-6_15

23. Pavluk N., Denisov A., Kodyakov A., Ronzhin A. Mechanical engineering of leg joints of anthropomorphic robot. *MATEC Web of Conf.*, vol. 77, p. 04006, EDP Sciences, 2016. doi:10.1051/matecconf/20167704006

24. *Determination of the center of pressure of Antares' foot during stride*. Available at: https://www.youtube.com/watch?v=NoYQ1oqr84M (accessed 30 May 2022).

25. *Center of pressure of Antares foot at pressure on each sensor cell*. Available at: https://www.youtube.com/watch?v=zRU3VnGsblk (accessed 30 May 2022).

26. Pitas I., Venetsanopoulos A. N. Nonlinear digital filters: Principles and applications. Ser. The Springer International Series in Engineering and Computer Science, vol. 84. Springer Science & Business Media, 2013. 392 p. doi:10.1007/978-1-4757-6017-0

27. *Simulation of the step of the ANTARES pedipulator*. Available at: https://www.youtube.com/watch?v=qlh-5NHOwMK0 (accessed 30 May 2022).

28. *Medium-density performance line ARM®-based 32-bit MCU*. Available at: https://www.st.com/resource/en/datasheet/stm32f103c8.pdf (accessed 30 May 2022).

29. Sardain P., Bessonnet G. Forces acting on a biped robot. Center of pressure — zero moment point. *IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans*, 2004, vol. 34, no. 5, pp. 630–637.

30. Galin R., Shiroky A., Magid E., Meshcheryakov R., Mamchenko M. Effective functioning of a mixed heterogeneous team in a collaborative robotic system. *Informatics and Automation*, 2021, vol. 20, no. 6, pp. 1224–1253 (In Russian). doi:10.15622/ia.20.6.2

УДК 681.586.2
doi:10.31799/1684-8853-2022-4-20-28

Обработка сигналов емкостных датчиков силы, установленных в стопу антропоморфного робота

К. Д. Крестовников[a], младший научный сотрудник, orcid.org/0000-0001-6303-0344
А. А. Ерашов[a], младший научный сотрудник, orcid.org/0000-0001-8003-3643, erashov.a@iias.spb.su
[a]Санкт-Петербургский Федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

Введение: основанное на компактных емкостных датчиках силы силомоментное очувствление функциональных поверхностей роботов позволяет существенно улучшить их взаимодействие с окружающей средой и людьми. Емкостные датчики силы обеспечивают высокую точность и скорость измерений, но электромагнитные помехи могут оказывать существенное влияние на их выходной сигнал. При последующей обработке сигналов должно учитываться влияние внешних помех, что увеличивает время вычислений. **Цель:** применить разработанную интерфейсную схему для обработки сигналов от емкостных первичных преобразователей силы в реальном роботе. **Результаты:** экспериментальная проверка разработанных решений заключалась в моделировании шага педипулятора антропоморфного робота с расчетом координат центральной точки давления на стопу с установленными четырьмя емкостными датчиками. Программная фильтрация и измерительные возможности микроконтроллера позволили добиться отношения сигнал/шум примерно 62,24 дБ, что дает возможность замкнутой системе автоматического регулирования корректно функционировать. Среднее время расчета координат центра давления на стопу при программной фильтрации сигнала на бортовом компьютере робота составило от 3,1 до 6,1 мс, что соответствует требованиям к сенсорной системе шагающего робота. **Практическая значимость:** интерфейсная схема имеет возможность масштабировать количество подключаемых первичных преобразователей силы, а программная обработка позволяет нормировать сигналы преобразователей путем применения рассчитанных поправочных коэффициентов. Предлагаемые решения могут быть использованы в различных робототехнических системах для измерения силы в режиме реального времени.

Ключевые слова — датчики силы, емкостные датчики, гуманоидные роботы, сенсорная система робота, цифровая обработка сигналов.

Для цитирования: Krestovnikov K. D., Erashov A. A. Signal processing of capacitive force sensors installed in the foot of an anthropomorphic robot. *Информационно-управляющие системы*, 2022, № 4, с. 20–28. doi:10.31799/1684-8853-2022-4-20-28
For citation: Krestovnikov K. D., Erashov A. A. Signal processing of capacitive force sensors installed in the foot of an anthropomorphic robot. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 20–28. doi:10.31799/1684-8853-2022-4-20-28

УДК 004.89
doi:10.31799/1684-8853-2022-4-29-43

Научные статьи
Articles

Исследование путей и способов повышения результативности выявления компьютерных атак на объекты критической информационной инфраструктуры

В. Н. Кузьмин[a], *доктор воен. наук, профессор, orcid.org/0000-0002-6411-4336, vka@mil.ru*
А. Б. Менисов[a], *канд. техн. наук, докторант, orcid.org/0000-0002-9955-2694*
[a]*Военно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ*

Введение: *в эпоху информационных технологий практически все организации сталкиваются с широким спектром автоматизированных и быстро распространяющихся угроз безопасности информации. Это обусловлено не только растущей сложностью, разнообразием и масштабом цифровизации, но и увеличением угроз и областей их возможной реализации.* ***Цель:*** *сравнить возможные пути повышения результативности подходов к выявлению компьютерных атак на объекты критической информационной инфраструктуры: обнаружение редкого события, аномалии и новизны функционирования объектов критической информационной инфраструктуры.* ***Результаты:*** *принцип работы предлагаемого (результативного) подхода к обнаружению компьютерных атак заключается в выявлении и отделении аномалий от нормального функционирования объектов с использованием концепции динамического изменения меток для переменного класса с течением времени. Динамическое обнаружение новизны сравнивается с другими подходами по показателю F1-меры. Для данных SWaT, который представляет собой макет объекта критической информационной инфраструктуры — автоматизированной системы управления, было определено, что количество выявленных атак с использованием предложенного подхода увеличилось на 7 %.* ***Практическая значимость:*** *результаты исследований показали снижение риска проведения (развития) компьютерной атаки на объектах критической информационной инфраструктуры. Возможное целевое применение подхода динамического обнаружения новизны заключается в оптимизации средств защиты информации на объектах критической информационной инфраструктуры, а также интеграции предложенного подхода в систему информационной безопасности как интеллектуального детектора.*

Ключевые слова *— информационная безопасность, технологии искусственного интеллекта, критическая информационная инфраструктура, нейронные сети.*

Для цитирования: Кузьмин В. Н., Менисов А. Б. Исследование путей и способов повышения результативности выявления компьютерных атак на объекты критической информационной инфраструктуры. *Информационно-управляющие системы*, 2022, № 4, с. 29–43. doi:10.31799/1684-8853-2022-4-29-43
For citation: Kuzmin V. N., Menisov A. B. A study of ways and solutions to increase the efficiency of detecting computer attacks on the objects of critical information infrastructure. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 29–43 (In Russian). doi:10.31799/1684-8853-2022-4-29-43

Введение

Современный мир переживает период бурного развития и трансформации. В этих условиях достижение цели обеспечения информационной безопасности объектов критической информационной инфраструктуры (КИИ; Critical Information Infrastructure, CII) может осуществляться путем реализации государственной политики, прежде всего в части обоснования и совершенствования комплекса мер, направленных на решение проблемы защиты объектов жизнеобеспечения населения, организаций оборонно-промышленного, атомного энергопромышленного, ядерного оружейного, химического, топливно-энергетического комплексов страны и объектов транспортной инфраструктуры от компьютерных атак.

Среди мер по обеспечению безопасного функционирования объектов КИИ выделяют регламентацию правил и процедур реагирования на компьютерные атаки, их выявление и анализ, а также защиту информации и информирование регуляторов, устранение последствий и принятие мер по недопущению повторного возникновения.

Критическая информационная инфраструктура — это сложная система, элементы которой используют различные программно-аппаратные компоненты при функционировании. Включение Интернета вещей (IoT) в КИИ открывает новые возможности злоумышленникам использовать уязвимости системы для проведения компьютерных атак [1].

За десятилетие, прошедшее с момента компьютерной атаки Stuxnet (2010 г.), увеличилось не только количество целевых компьютерных атак (далее — атак) на объекты КИИ, но и их разнообразие, а также усложнились технологии проведения [2]. На основе выполненного анализа в табл. 1 и 2 отображены действия злоумышленников известных целевых атак [3–12] с выделением соответствующих последствий.

■ ***Таблица 1.*** Сравнение атак на объекты КИИ
■ ***Table 1.*** Comparison of attacks on CII objects

Действие злоумышленников	Атака на объекты КИИ						
	Stuxnet	Havex	Black Energy	Немецкий завод	Duqu2.0	Crash Override	TRISIS
Доступ в Интернет	×	√	√	√	√	√	?
Аудит сети	√	√	√	?	√	√	?
Несанкционированный доступ	√	√	√	?	√	√	√
Отображение производственного процесса	√	√	×	?	×	√	√
Взломанный хост	√	√	√	√	√	√	√
Фальсифицированные данные	√	×	×	?	×	√	×
Несанкционированное выполнение программ	√	√	√	?	√	√	√
Сбор учетных данных	×	×	√	√	×	√	×
Вредоносная прошивка	√	×	×	?	×	√	√

× — не свойственно для атаки;
√ — действия, характерные для атаки;
? — данные отсутствуют.

■ ***Таблица 2.*** Последствия проведения атак на объекты КИИ
■ ***Table 2.*** Consequences of attacks on CII objects

Действие злоумышленников	Последствие
Доступ в Интернет	1. Возможность получать данные из сети через хосты с подключением к Интернету 2. Возможность загружать программное обеспечение
Аудит сети	1. Анализ хостов в сети для будущих атак 2. Предоставление информации о сети
Несанкционированный доступ	1. Утечка данных 2. Загрузка дополнительных исполняемых файлов
Отображение производственного процесса	1. Предоставление информации о цели 2. Возможность разработать целевое вредоносное программное обеспечение для этого процесса
Взломанный хост-компьютер	1. Предоставление информации о системе. Злоумышленники могут создавать профили пользователей системы и определять шаблоны, которые можно использовать для дальнейшего проникновения в сеть 2. Идентификация активов посредством анализа журналов и мониторинга сети
Фальсифицированные данные	1. Потеря доверия к сети, устройству или программному обеспечению 2. Изменение данных 3. Физический ущерб
Несанкционированное выполнение программ	Предоставление дополнительных возможностей по действию в системе и позволение в ней закрепиться
Сбор учетных данных	Предоставление подлинного, замаскированного доступа к службам
Вредоносная прошивка	1. Возвращение оператору фальсифицированных и недостоверных сведений 2. Физический урон

Перечисленные действия злоумышленников вызывают негативные последствия для объектов КИИ. Так, устройства, подключенные к Интернету, позволяют проводить удаленное взаимодействие, хищение данных и загрузку дополнительного программного обеспечения. Каждая из проанализированных атак манипулировала устройством, подключенным к Интернету, за исключением Stuxnet, которая была способна выполнять сложные автоматизированные действия для достижения своих целей.

Несанкционированный доступ к сети — единственный общий элемент для всех атак — заключается в возможности взаимодействовать с другими сетевыми устройствами для получения дополнительных привилегий доступа или дополнительной информации о цели. Обычно несанкционированный доступ выполняется, когда противник уже проник в систему. Важным шагом является снятие образа устройств после их обнаружения. Например, Stuxnet и TRISIS идентифицировали целевые устройства по результатам несанкционированного сканирования сети.

В большинстве атак злоумышленники демонстрируют реализацию угроз, которую можно разделить на следующие этапы: сбор информации, получение первоначального доступа, внедрение и использование вредоносного кода, закрепление в системе и сети, управление вредоносным кодом и компонентом, повышение привилегий, сокрытие действий, получение доступа к другим компонентам, сбор и вывод информации и неправомерный доступ или воздействие.

Существует достаточное количество исследований по определению возникновения угроз безопасности информации с помощью традиционных средств [13–16]. Стоит указать, что ведущие организации в области информационной безопасности [17] отмечают растущее использование возможностей технологий искусственного интеллекта в текущем ландшафте угроз:

— расширение существующих угроз, которое связано со сложными компьютерными атаками на большое количество потенциальных целей и низкой стоимостью атак;

— введение новых угроз, связанных с задачами, которые были бы невыполнимы для человека;

— изменение типичного характера угроз, которое включает в себя новые атрибуты автоматизированных, высокоэффективных, трудно определяемых и крупномасштабных атак в ландшафте угроз.

Для описания способов реализации угроз наиболее хорошо зарекомендовали себя следующие подходы к представлению угроз безопасности информации:

— деревья атак [18], предложенные в 1999 г. Брюсом Шнайером;

— цепочки Kill Chain [19], разработанные в 2011 г. компанией Lockheed Martin для обнаружения нарушителей на протяжении всего жизненного цикла компьютерной атаки;

— набор тактик и техник поведения нарушителей MITRE ATT & CK [20];

— модель жизненного цикла компьютерной атаки [21], в которой особое внимание уделяется моделированию APT-атак и демонстрируется повторяющийся характер нарушителей для дальнейшего повышения привилегий.

Таким образом, система защиты имеет исходные данные о предыдущих состояниях объекта КИИ, а для изменения состояния можно использовать следующую модель, в которой выделены основные параметры, влияющие на функционирование объектов КИИ:

$$\overline{f_k} = \overline{f_k}\left\{s^a_{i_1};\ s^{\tilde{a}}_{i_2};\ s^p_{i_3};\ s^{\tilde{p}}_{i_4};\ g_{\omega_k};\ t_k;\ d_j;\ l_s;\ z_g;\ r_k\right\},$$

где k — номер конкретного объекта КИИ, $k = 1, \ldots, K$; i — номер признака, $i = 1, \ldots, N$, т. е. результата действия злоумышленника; $s^a_{i_1}$, $s^a_{i_2}$ — данные наблюдения предыдущих состояний объекта КИИ, $i_1 = 1, \ldots, m_1$, $i_2 = m_1 + 1, m_2$; $s^p_{i_3}$, $s^{\tilde{p}}_{i_4}$ — данные текущего контроля объекта КИИ, $i_3 = m_2 + 1, m_3$, $i_4 = m_3 + 1, m_4 = m$; $\sim$ — знак, указывающий на признаки, которые подвержены влиянию данной совокупности внешних условий; g_{ω_k} — совокупность внешних условий, $\omega_k = 1, \ldots, \Omega_k$; t_k — условная координата времени, показывающая полноту информации о k-м объекте КИИ; d_j — класс атаки на объект КИИ, $j = 1, \ldots, N$; l_s — методы защиты объекта КИИ, $s = 1, \ldots, S$; z_g — последующие состояния k-го объекта КИИ, $g = 1, \ldots, G$; r_k — помеха, искажающая действительное состояние k-го объекта КИИ.

Формализованная постановка задачи исследования

Процесс обеспечения безопасности объектов КИИ является сложным циклическим процессом, включающим в себя сбор и обработку данных различных систем, определение состояния объекта КИИ, выбор стратегии защиты и проведение защитных мер. Управление процессом обеспечения безопасности функционирования объектов КИИ может происходить с участием и без участия персонала.

Процесс обеспечения безопасности может состоять из различного числа циклов — от двух до многих десятков. Поэтому в каждом цикле проводится оценка и коррекция управляющего воздействия. Сущность обеспечения защиты

объектов КИИ состоит в следующем. После того как установлены специфические особенности функционирования, можно переходить к построению прогноза и исхода защиты объекта КИИ.

Пусть известна модель защищаемого объекта КИИ $\bar{f}_k$. Пусть l_s – защитные воздействия (методы), которые могут объект КИИ $\bar{f}_k$ из состояния z_{g_i} перевести в новое состояние $z_{g'}(g, g', g'' = 0, ..., G; g \neq g' \neq g'')$, и пусть $p(z_{g_j} \to z_{g'})_{l_s}$ – вероятность такого перехода. Обозначим через e_{s_j} меру результативности l_s-го мероприятия защиты объекта КИИ. Тогда задачу нахождения оптимальной совокупности защитных мер можно сформулировать следующим образом.

Необходимо найти такую совокупность защитных воздействий l_s^*, чтобы мера их результативности была максимальна. В этом случае мера результативности

$$e_{s_j}(z_{g_j} \to z_{g'})_{l_s} = p(d_j)_{z_g} p(z_{g_j} \to z_{g'})_{l_s},$$

где $p(d_j)_{z_g}$ – вероятность проведения атаки, а максимальное значение меры результативности, или оптимальная защита объекта КИИ для $(z_{g_j} \to z_{g'})_{l_s}$, достигается при

$$e^*_{s_j} = \min_{l_s} e_{s_j}(z_{g_j} \to z_{g'})_{l_s}.$$

Это определение действует для тех пар $(z_{g_j} \to z_{g'})_{l_s}$, для которых справедливо утверждение о том, что $z_{g'}$ лучше, чем z_{g_j}.

Предположим, что общая результативность защитных мер e_j является аддитивной функцией, состоящей из e_{s_j}. Тогда для определения оптимальной совокупности защитных воздействий можно сформулировать ограничение исследования: оптимальная совокупность защитных воздействий обладает тем свойством, что, каково бы ни было первоначальное защитное воздействие l_s при состоянии защищенности объекта КИИ z_{g_j}, последующие защитные воздействия должны быть оптимальны относительно первоначального защитного воздействия. Исходя из этого принципа максимальную результативность защитных воздействий можно получить в следующем виде:

$$e_j = \min_{l_s} \sum | e_{s_j}(z_{g_j} \to z_{g'})_{l_s} + e_{s_j}(z_{g'} \to z_{g''})_{l_s} |.$$

Таким образом, выбор оптимальной совокупности защитных воздействий начинается с выявления нарушения функционирования объектов (атак на объекты) КИИ.

Пути повышения результативности выявления компьютерных атак на объекты КИИ

В то время как проблемы безопасности объектов КИИ активно рассматриваются в научных кругах и ИТ-сообществе, все решения являются ограниченными для различных условий. Атаки на объекты КИИ в основном остаются слабо идентифицируемыми для традиционных средств информационной безопасности, таких как системы обнаружения вторжений (Intrusion Detection System, IDS) и антивирусные программы. Кроме того, протоколы, используемые системами контроля объектов КИИ (например, Modbus [22] или DNP3 [23] и стандарты IEC [24]), не защищены должным образом традиционными IDS. Следовательно, для защиты объектов КИИ необходимо разрабатывать надежные механизмы безопасности.

В литературе использовались различные подходы для разработки IDS, в том числе на основе методов машинного обучения [25–27]. Большинство этих методов используют доступные данные для разработки модели, которая демонстрирует нормальное поведение объекта КИИ, а затем идентифицирует все различные варианты поведения как ненормальные. Поскольку эти модели обучены только конкретным типам атак, они не могут обнаруживать скрытые или новые типы атак [28]. В последние годы результаты исследований в различных областях внесли свой вклад в решение ряда связанных проблем, основными сущностями которых являются редкие события, аномалии и новизна.

Таким образом, в качестве путей повышения результативности выявления компьютерных атак на объекты КИИ будем рассматривать обнаружение редкого события, аномалии функционирования и новизны функционирования.

Обнаружение редкого события

Почти все работы, в которых используется термин *редкое событие*, представляют время наблюдения за характеристиками функционирования объектов КИИ как общую характеристику. То есть данные могут быть разделены по часовым интервалам или другим фиксированным значениям.

В ранее описанных задачах цель состоит в предсказании наступления редкого события за ограниченный промежуток времени. Основной характеристикой обучения модели выявления атак на объекты КИИ как редкого события с точки зрения классификации является то, что функционирование объекта КИИ представляет собой временной ряд. Цель состоит в том, чтобы классифицировать новые данные о состоянии как

редкие (R, когда произошло редкое событие) или нормальные (N, когда не произошло никакого события), используя ранее обученную модель. Этот подход известен в машинном обучении как классификация временных рядов [29].

Вместе с тем из-за временной природы проблемы в литературе можно найти два различных подхода к обнаружению редких событий.

В первом подходе рассматривается полноразмерная классификация временных рядов. Например, в работе [30] представлен подход к определению неисправности жесткого диска в течение фиксированного периода времени. Авторы формируют модель машинного обучения, используя записанные на жестком диске сенсорные измерения в разное время. Затем, учитывая новые данные датчиков жесткого диска, обнаруживают сбой — редкое событие.

Другой подход к определению редких событий заключается в том, чтобы классифицировать новые наблюдения (временные ряды) как можно раньше, предпочтительно до того, как будет доступен полный временной ряд. Этот подход известен как ранняя классификация временных рядов.

Общий процесс классификации представлен на рис. 1.

В большинстве задач, направленных на определение состояния функционирования объектов КИИ как редкое событие, стоит отметить несбалансированное распределение классов. Таким образом, классификация редких событий может быть формализована как задача классификации несбалансированных временных рядов. В частном случае редкое событие определено как $P(R) << P(N)$.

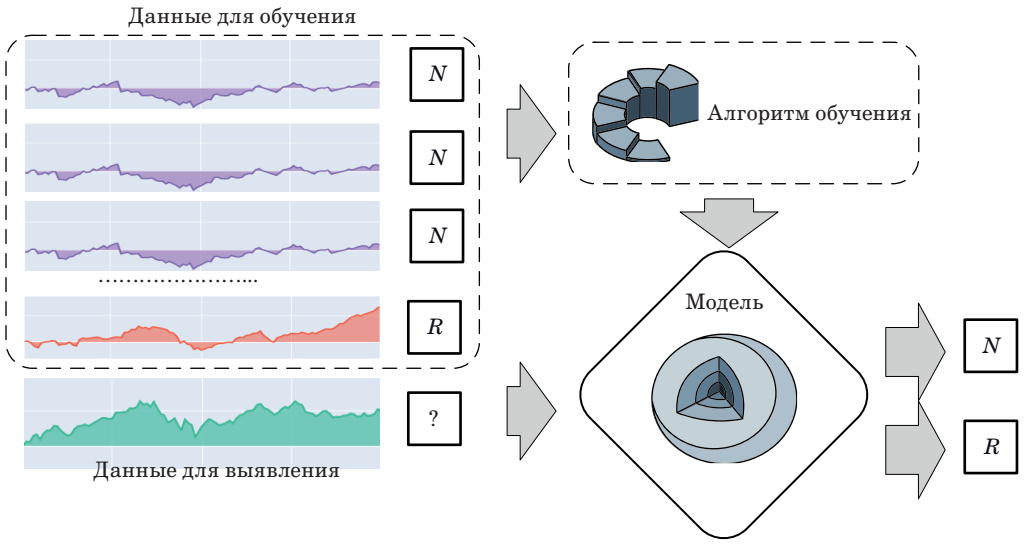
Временной ряд — это упорядоченная совокупность результатов наблюдений пар (временная метка, значение) фиксированной длины m:

$$z_g = \left\{\left(t_1, \tilde{s_1^p}\right), \ldots, \left(t_i, \tilde{s_i^p}\right), \ldots, \left(t_m, \tilde{s_m^p}\right)\right\},$$

где t — условная координата времени, показывающая полноту информации об объекте КИИ, а $\tilde{s_1^p}$ — характеристика функционирования объекта КИИ.

Для выявления редкого события необходим набор для обучения $\{(z_{g1}, y_1), \ldots, (z_{gn}, y_n)\}$. Цель — построить модель-классификатор, которая способна предсказать метку класса новых временных рядов.

Однако несмотря на то, что выявление редкого события в качестве ключевого компонента рассматривает оценку состояния в определенный момент времени, в некоторых решениях данные преобразуются без учета этой характеристики. Например, в работе [31] данные состоят из нескольких измерений датчиков жесткого диска с различными временными интервалами. Поэтому для одного и того же устройства доступно множество показаний одних и тех же датчиков. Однако авторы не учитывали порядок, в котором были записаны измерения, и, учитывая новые неупорядоченные измерения на жестком диске, модель классифицировала диск как неисправный или нормальный. Следовательно, временная природа данных использовалась не совсем корректно. Таким образом, изложенный подход рассматривает проблему как задачу нестационарной несбалансированной классификации, аналогичную тем, которые встречаются в решениях обнаружения аномалий.



■ ***Рис. 1.*** Представление задачи классификации временных рядов
■ ***Fig. 1.*** Representation of the task of classifying time series

Обнаружение аномалий

Большинство проблем, описывающих поиск аномалий (A), не имеют временной природы (рис. 2).

В алгоритме обучения обнаружению аномалий аномальные экземпляры данных редки из-за несбалансированного распределения между нормальными и аномальными классами [32]. Поэтому проблема может быть формально описана как несбалансированная классификация. Что касается распределения вероятностей аномального класса, то $P(A) << P(N)$.

Формально функционирование объекта КИИ определяется как $s^{\tilde{p}} = \left(s^{\tilde{p}}_{i_1}, ..., s^{\tilde{p}}_{i_m}\right)$. Учитывая, что существует набор данных $z_g = \left\{\left(s^{\tilde{p}}_{i_1}, y_1\right), ..., \left(s^{\tilde{p}}_{i_m}, y_m\right)\right\}$, в котором y представляет метку класса (нормальное функционирование или аномалия) соответствующего экземпляра данных, необходимо обучить модель-классификатор, способный предсказать метку класса любого нового экземпляра как можно точнее.

Обнаружение новизны

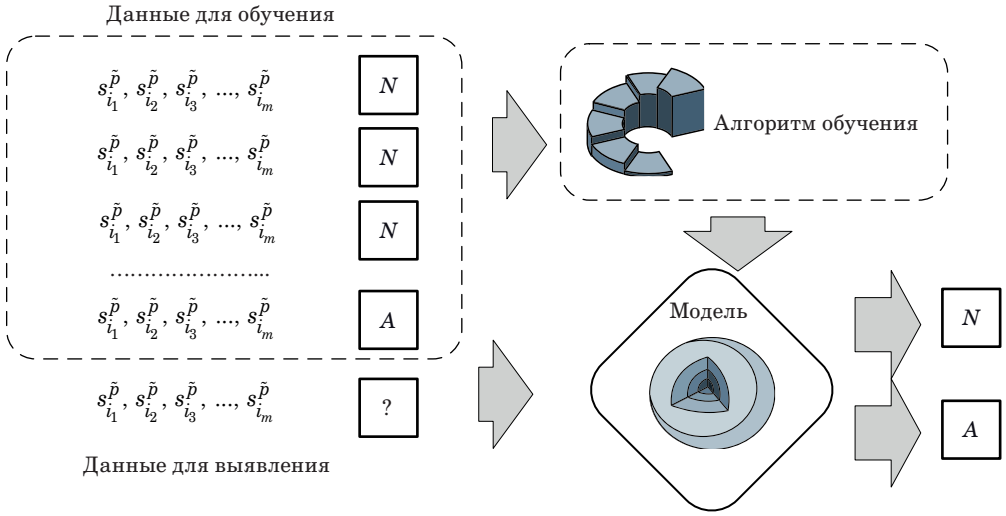
В большинстве работ, направленных на поиск новизны, модель обучают с использованием набора данных, содержащего только один класс. Например, в работах [33, 34] действия системы классифицируются IDS как новые или нормальные. Новый экземпляр соответствует неподдерживаемому или неожиданному действию системы. Для обучения модели используются только обычные действия в системе, собранные в защищенной среде. Когда наступает новое (неизвестное) действие в системе, модель-классификатор предсказывает его как обычное или новое.

Существует два различных способа обучения модели — статическое и динамическое.

Статическое обнаружение новизны может быть представлено как задача бинарной классификации. При наличии набора данных, состоящего только из одного класса, обучается модель. Эта модель изучает границу принятия решения, которая изолирует нормальное поведение. При появлении неизвестного экземпляра его классифицируют как новый или как нормальный. В этих рамках усилия сосредоточены на классификации нормального класса. Поэтому для оценки результативности таких моделей-классификаторов обычно максимизируется полнота определения нормального класса. Формально обучающий набор генерируется только из $P\left(s^{\tilde{p}} \mid z_g = N\right)$. На этапе обучения, даже если классификатор обучается с использованием информации только об одном классе (нормальном функционировании объекта КИИ), он строится с учетом того, что существует другое поведение, отличное от нормального.

Формально любое функционирование объекта КИИ определяется как $s^{\tilde{p}} = \left(s^{\tilde{p}}_{i_1}, ..., s^{\tilde{p}}_{i_m}\right)$. С учетом размеченного набора данных для обучения $z_g = \left\{\left(s^{\tilde{p}}_{i_1}, y_1 = N\right), ..., \left(s^{\tilde{p}}_{i_m}, y_m = N\right)\right\}$. Цель заключается в том, чтобы обучить модель-классификатор, которая будет в состоянии предсказать разницу между нормальным функционированием объекта КИИ и новым.

Динамическое обнаружение новизны (*обнаружение эволюционирующих классов, будущих*



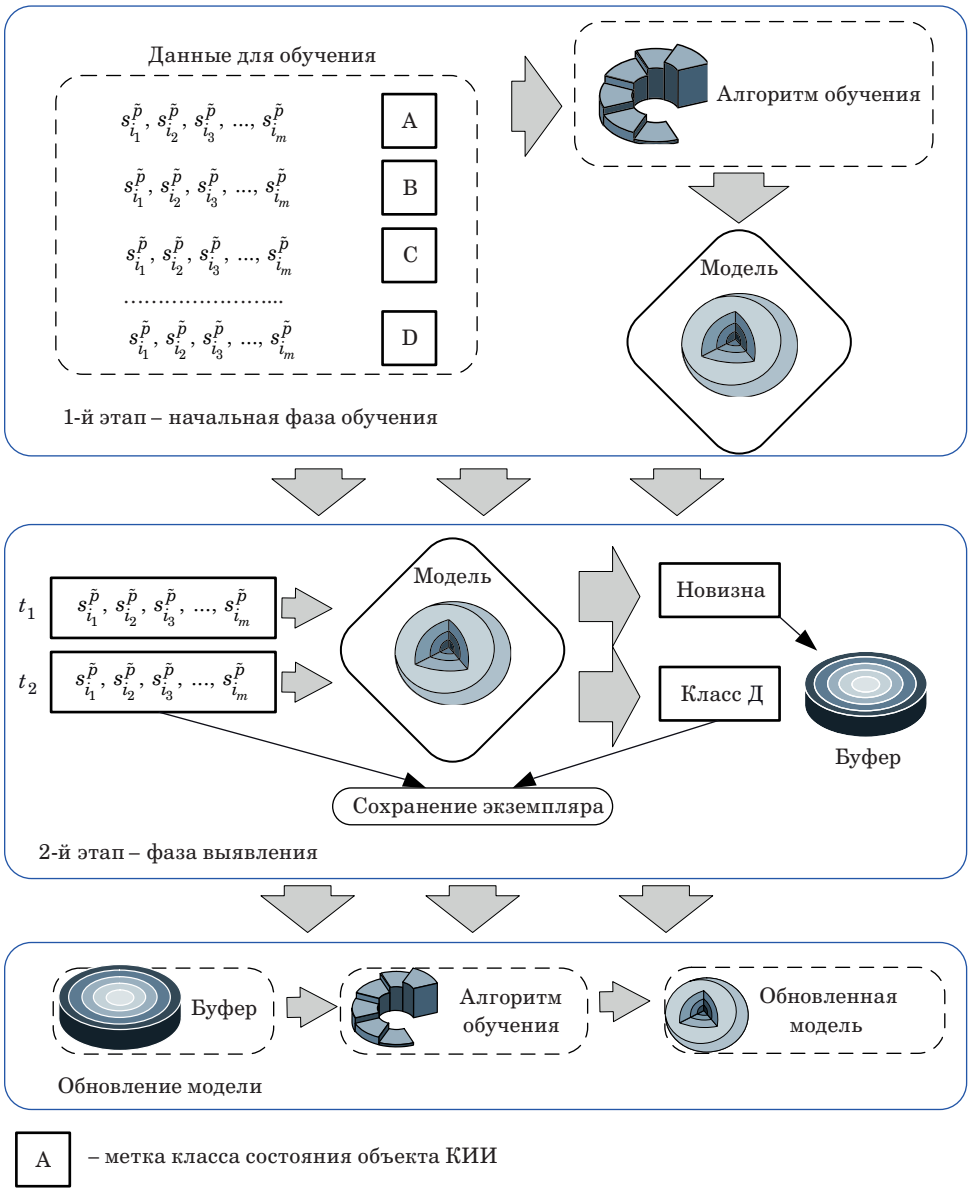
■ ***Рис. 2.*** Представление задачи выявления аномалий
■ ***Fig. 2.*** Representation of the anomaly detection task

классов, *новых классов*). Этот способ обучения может быть формализован как задача классификации, в которой количество меток для переменной класса неизвестно. Другими словами, распределение вероятностей z_g динамически изменяется в процессе классификации, поэтому классификатор должен адаптироваться к этим изменениям. Когда появляется новый экземпляр, модель должна классифицировать его среди текущих классов или хранить в буфере. Учитывая жизненный цикл функционирования объектов КИИ и меняющийся ландшафт угроз, классы состояний объекта КИИ z_g могут появляться, удаляться и появляться вновь. Следовательно, классификатор должен быть обновлен для этих изменений с учетом того, что время адаптации имеет особое значение.

Способ обучения динамическому обнаружению новизны можно разделить на два этапа (рис. 3).

1-й этап — *начальная фаза обучения*. С учетом помеченного обучающего набора данных с набором меток $\left(d_j\right)_{z_g}$ строится модель-классификатор при состоянии объекта КИИ z_{gj}.

2-й этап — *фаза выявления*. Новые классы атак и других воздействий на объекты КИИ могут появляться и исчезать, а старые классы могут изменяться. Данный этап может быть описан как стадия подготовки к адаптации модели-классификатора для обработки потока данных (бес-



■ ***Рис. 3.*** Динамическое обнаружение новизны
■ ***Fig. 3.*** Dynamic novelty detection

■ ***Таблица 3.*** Обобщенные основные характеристики путей повышения результативности выявления атак на объекты КИИ

■ ***Table 3.*** Generalized main characteristics of ways to improve the effectiveness of detecting attacks on CII objects

№ п/п	Путь повышения результативности	Методы	Основные характеристики
1	Обнаружение редкого события	Логистическая регрессия [35, 36], расстояние Кульбака — Лейблера [37], нейронные сети [29], байесовские сети [38], метод опорных векторов [39]	Временные ряды Несбалансированная классификация Все классы в обучающих данных
2	Обнаружение аномалии	Автоэнкодеры [40], k-средних [41], нейронные сети [42–45]	Статистические данные Несбалансированная классификация Все классы в обучающих данных
3	Обнаружение новизны (статическое)	Нейронные сети [46, 47], изолирующий лес [48]	Один класс в обучающих данных
4	Обнаружение новизны (динамическое)	Нейронные сети [46], автоэнкодеры [48]	В обучающих данных не все возможные классы

конечная последовательность результатов наблюдений). В момент времени t_k текущий классификатор предсказывает новый вариант функционирования объекта КИИ. Если новое состояние объекта КИИ не может быть классифицировано в текущем наборе, оно хранится отдельно в буфере и модель не изменяется. Как только буфер заполняется, классификатор обновляется и набор меток t_{k+n} изменяется.

Таким образом, динамическое обнаружение новизны обладает возможностью адаптации к изменениям состояния функционирования объектов КИИ. Характеристики рассмотренных путей повышения результативности выявления атак на объекты КИИ представлены в табл. 3.

Для методов, представленных в табл. 3 (пп. 1 и 2), характерен основной недостаток — зависимость от опыта экспертов по маркировке (разметке) данных. Еще одним недостатком этих подходов является то, что новые аномалии, которые не были частью обучающих данных, могут не обнаруживаться.

Подходы п. 3 могут обнаруживать только кратковременные отказы, вызванные компьютерными атаками [49, 50]. Кроме того, обнаружение новизны должно быть в потоковом режиме. Это позволяет системным администраторам вмешиваться в текущую атаку или устранять проблемы с производительностью системы. В связи с этим подходы статического обнаружения новизны, учитывающие данные устаревших событий, не подходят для современных систем обнаружения компьютерных атак.

Стоит отметить, что для методов п. 4 (динамического обнаружения новизны) не характерно предположение, что данные функционирования объектов КИИ стационарны. При этом всегда существует риск неадекватных защитных мер в будущем из-за «неучтенной динамики». Альтернативные подходы основаны на онлайн-обучении моделей [51]. Однако такие решения обладают угрозами безопасности информации [52], при реализации которых злоумышленники могут снизить эффективность этих решений.

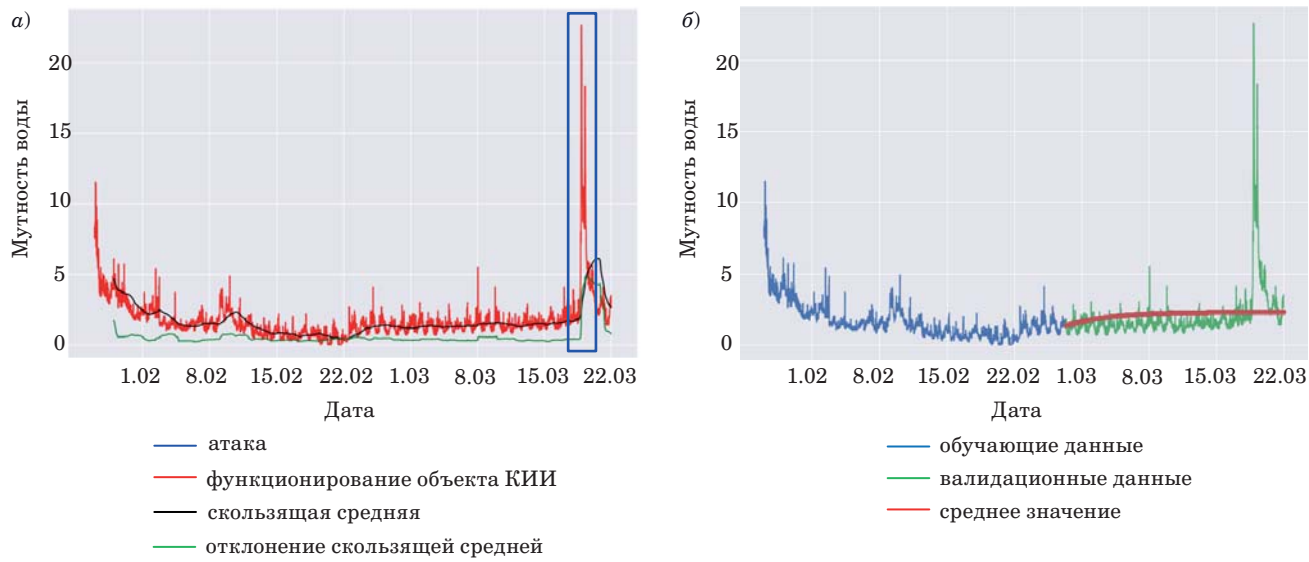
Эксперимент

Для выбора из рассмотренных путей повышения результативности выявления атак на объекты КИИ был проведен эксперимент. Целью эксперимента была проверка гипотезы, что не оптимизированные алгоритмы машинного обучения будут иметь разный результат и могут быть проранжированы по метрикам качества.

Данные

Отсутствие надежных и общедоступных наборов данных функционирования объектов КИИ является фундаментальной проблемой для исследователей, изучающих защищенность таких объектов от различного типа атак и других воздействий. Персонал реальных объектов КИИ не может предоставлять точные наборы данных, поскольку ошибки или атаки в лучшем случае можно только предполагать.

Для проведения эксперимента вместо объекта КИИ были использованы данные испытательного стенда системы очистки воды (SWaT) [53]. SWaT представляет собой уменьшенную версию реальной городской водоочистной станции, производящей 25 литров воды в минуту с помощью мембранных установок и обратного осмоса. Стенд функционировал в двух состояниях: нормальном и при атаке на информационную инфраструктуру.



■ ***Рис. 4.*** Пример функционирования системы очистки воды (SWaT) после атаки: *а* — атака; *б* — разделение на выборки данных

■ ***Fig. 4.*** Example of the functioning of the water treatment system (SWaT) after attack: *а* — attack; *б* — data sampling

Для атаки на систему очистки воды использовался системный подход. Входом атаки может быть физический элемент или коммуникационная сеть, соединяющая датчики или исполнительные механизмы с контроллерами и системой SCADA. SWaT состоит из шести уровней, каждый из которых содержит разное количество датчиков и исполнительных механизмов. Набор данных содержит сетевой трафик и все значения, полученные от 51 датчика и исполнительных механизмов в течение 11 дней непрерывной работы (семь дней в обычном режиме и четыре дня с 41 атакой). Данные помечены в соответствии с нормальным и аномальным функционированием.

Результирующее действие на функционирование — качество очистки воды — представлено на рис. 4, *а*. Набор данных содержит физические характеристики, связанные с процессом водоподготовки, а также сетевой трафик на испытательном стенде. Для проведения эксперимента данные были разделены на обучающую и валидационные части (рис. 4, *б*), содержащие результаты проведения атак.

Оценивание качества выявления нарушения функционирования объектов КИИ

Для оценки эффективности различных подходов из-за (сильно) несбалансированного распределения классов общие показатели качества, такие как точность, недостаточно информативны. Поэтому метриками качества для сравнения было принято использовать метрики пропуска и ошибки выявления нового (аномального) функционирования объекта КИИ, а также F-меру.

F-мера была рассчитана путем рассмотрения двух других метрик: полноты и точности. Все показатели перечислены в следующих уравнениях.

Показатель точности

$$FE = \frac{z_{g'} \to z_g}{z_g + z_{g'} \to z_g},$$

где z_g — количество экземпляров, правильно классифицированных как старый класс; $z_{g'} \to z_g$ — количество экземпляров из нового класса, классифицированных как старый класс.

Далее оценили метрику полноты, которая представляет количество правильно определенных объектов из общего числа объектов в наборе данных.

Показатель пропуска нового (аномального) функционирования объекта КИИ

$$MN = \frac{z_g \to z_{g'}}{z_{g'} + (z_g \to z_{g'})},$$

где $z_{g'}$ — количество экземпляров, правильно классифицированных как новый класс; $z_g \to z_{g'}$ — количество экземпляров из старого класса, классифицированных как новый класс.

Наконец, F-мера может быть вычислена на основе полученных значений:

$$F_1 = 2\frac{MN \cdot FE}{MN + FE}.$$

Валидации эксперимента

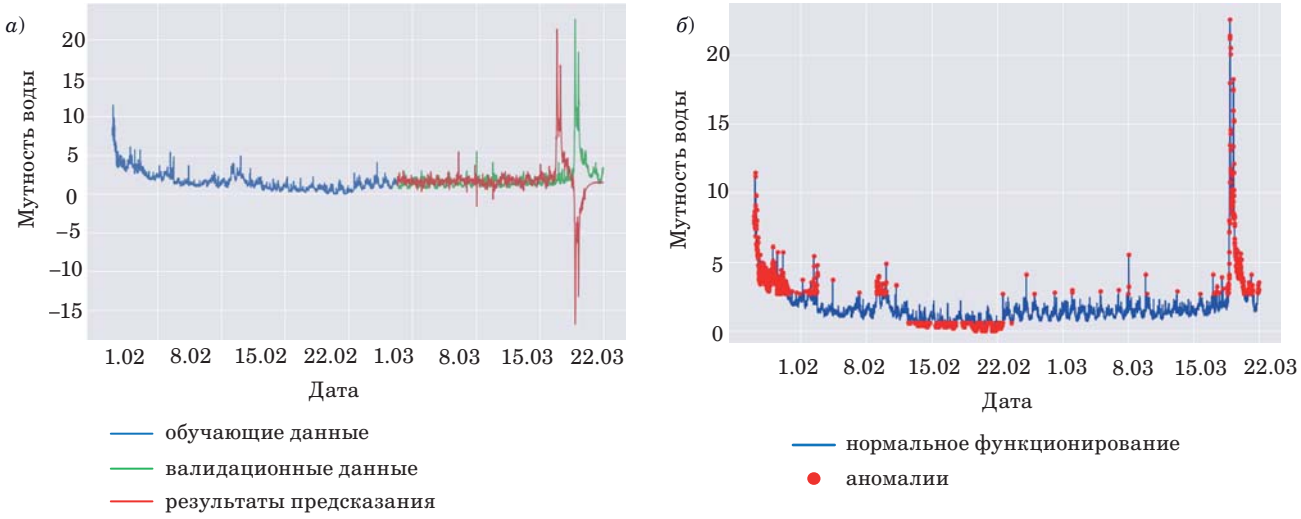
Результаты эксперимента представлены на рис. 5 и в табл. 4. Стоит отметить, что для статистического выявления новизны характерно отклонение результатов по времени (рис. 5, *а*), а для выявления редкого события (рис. 5, *б*) — ложноположительные результаты и низкая точность выявления начальной фазы атаки.

Проведенный эксперимент показал, что из-за ограниченного времени выявить можно только ограниченное количество атак, поэтому качество их выявления для любых алгоритмов должно быть крайне высоким. При динамическом выявлении качество (F-мера) повышается. Связано это с адаптацией под различные изменения функционирования из-за 41 атаки.

Стоит отметить, что для детекторов (автоэнкодеров и нейронных сетей) характерна зависимость от длины данных, взятых для обучения. Лучшие результаты были получены при размере пакетов 22 записи. Качество обнаружения сильно различается для разных размеров пакета, даже при небольших изменениях (21 или 23 записи). В целом оба детектора дают многообещающие результаты.

Заключение

В статье представлены основные данные сравнения путей повышения результативности выявления атак на объекты КИИ: редкого собы-



■ ***Рис. 5.*** Результаты выявления атак на объект КИИ: *а* — статистическое выявление новизны; *б* — выявление редкого события

■ ***Fig. 5.*** Representation of the results of detecting an attacks on the CII object: *а* — statistical novelty identification; *б* — identification of a rare event

■ ***Таблица 4.*** Результаты сравнения алгоритмов

■ ***Table 4.*** Algorithm comparison results

Редкое событие		Аномалия		Новизна (статическое выявление)		Новизна (динамическое выявление)	
Подход	F-мера	Подход	F-мера	Подход	F-мера	Подход	F-мера
Идеальный детектор	1	Идеальный детектор	1	Идеальный детектор	1	Идеальный детектор	1
Нейронные сети	0,8446	Автоэнкодеры	0,8301	Нейронные сети	0,8598	Автоэнкодеры	0,9045
Байесовские сети	0,8275	*k*-средних	0,8046	Случайный лес	0,8234	Нейронные сети	0,8687
Логистическая регрессия	0,7657	Нейронные сети	0,7945	Логистическая регрессия	0,7963	Логистическая регрессия	0,8174
Нулевой детектор	0	Нулевой детектор	0	Нулевой детектор	0	Нулевой детектор	0

тия, аномалии и новизны. Для этого были рассмотрены характеристики процесса выявления, исходных данных и наиболее репрезентативные методы.

Выявление известных сигнатур компьютерных атак по-прежнему критически важно, но оно не обеспечивает достаточной защищенности в ландшафте угроз безопасности КИИ. Данная технологическая отрасль должна избегать проблемы с изменяющимися атаками с помощью решений последнего поколения, применяя новые подходы, которые устраняют пробелы в разработке и функционировании сервисов и систем защиты.

Принцип работы предлагаемого подхода заключается в выявлении и отделении аномалий от нормальных наблюдений с использованием концепции динамического изменения меток для переменного класса с течением времени. Динамическое обнаружение новизны сравнили с другими подходами по показателю F1-меры. Для данных SWaT, который представляет собой макет объекта КИИ — автоматизированной системы управления, было установлено, что выявление атак с использованием динамического обнаружения новизны улучшилось на 7 %.

Дальнейшие исследования, на наш взгляд, должны быть направлены на:

1) повышение производительности подхода к обнаружению атак посредством сокращения признакового пространства функционирования объектов КИИ;

2) расширение предлагаемого подхода на гибридную систему обнаружения, использующую как данные процесса, так и данные сетевого трафика системы, чтобы улучшить качество обнаружения скрытых атак.

Финансовая поддержка

Работа выполнена в рамках гранта Президента Российской Федерации для государственной поддержки молодых российских ученых — кандидатов наук МК-2485.2022.4.

Литература

1. **Petrenko S. A.** Cyber resilient platform for Internet of Things (IIoT/IoT) ed systems: survey of architecture patterns. *Вопросы кибербезопасности*, 2021, № 2, с. 81–91. doi:10.21681/2311-3456-2021-2-81-91
2. **Maynard P., McLaughlin K., Sezer S.** Decomposition and sequential-AND analysis of known cyber-attacks on critical infrastructure control systems. *Journal of Cybersecurit*, 2020, vol. 6, iss. 1, pp. 1–20. doi:10.1093/cybsec/tyaa020
3. **Langner R.** Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 2011, vol. 6, iss. 3, pp. 49–51. doi:10.1109/MSP.2011.67
4. **Miller B., Rowe D.** A survey SCADA of and critical infrastructure incidents. *Proc. of the 1st Annual Conf. on Research in Information Technology*, New York, Oct. 2012, NY, United States, 2012, pp. 51–56. doi:10.1145/2380790.2380805
5. **Assante M. J., Lee R. M.** *The Industrial Control System Cyber Kill Chain*. SANS Institute InfoSec Reading Room, 2015. Vol. 1. 24 p.
6. **Rrushi J., Farhangi H., Howey C., Carmichael K., Dabell J.** A quantitative evaluation of the target selection of havex ics malware plugin. *Industrial Control System Security (ICSS) Workshop*, Los Angeles, California, USA, 2015, pp. 1–5.
7. *ICS-CERT A. Ongoing sophisticated malware campaign compromising ICS*. http://www.ics-cert.uscert.gov/alerts/ICS-ALERT-14-281-01B (дата обращения: 12.01.2022).
8. **Rrushi J. L.** Multi-range Decoy I/O defense of electrical substations against industrial control system malware. *Resilience of Cyber-Physical Systems: From Risk Modelling to Threat Counteraction*. F. Flammini (eds). Springer, Cham., 2019. Pp. 151–175. doi:10.1007/978-3-319-95597-1_7
9. **Lee R. M., Assante M. J., Conway T.** German steel mill cyber attack. *Industrial Control Systems*, 2014, vol. 30, iss. 62, pp. 1–15.
10. **Bencsáth B., Pék G., Buttyán L., Félegyházi M.** *Duqu: A Stuxnet-like malware found in the wild*. CrySyS Lab Technical Report, 2011. Vol. 14. 60 p.
11. **Zetter K.** *A cyberattack has caused confirmed physical damage for the second time ever*. https://www.wired.com/2015/01/german-steel-mill-hack-destruction/ (дата обращения: 12.01.2022).
12. **Alladi T., Chamola V., Zeadally S.** Industrial control systems: Cyberattack trends and countermeasures. *Computer Communications*, 2020, vol. 155, pp. 1–8. doi:10.1016/j.comcom.2020.03.007
13. **Саенко И. Б., Лаута О. С., Карпов М. А., Крибель А. М.** Модель угроз ресурсам ИТКС как ключевому активу критически важного объекта инфраструктуры. *Электросвязь*, 2021, № 1, с. 36–44. doi:10.34832/ELSV.2021.14.1.004
14. **Рыбкина О. В.** Построение модели угроз безопасности информации на основе математической модели Ланкастера. *Научно-техническое и экономическое сотрудничество стран АТР в XXI веке: тр. Всерос. науч.-практ. конф.*, Хабаровск, 20–23 апреля 2021 г., 2021, т. 1, с. 252–257.
15. **Бражук А. И.** Методика моделирования угроз компьютерных систем на основе предметно-ориентированных моделей. *Всерос. науч. конф. «Информационные технологии в моделировании и управле-*

нии: подходы, методы, решения», Тольятти, 20–22 апреля 2021 г., 2021, с. 94–101.

16. **Суханов И. Д., Рыбкина О. В.** Новые подходы к моделированию угроз безопасности информации. *Научно-техническое и экономическое сотрудничество стран АТР в XXI веке: тр. Всерос. науч.-практ. конф.*, Хабаровск, 20–23 апреля 2021 г., 2021, т. 1, с 277–282.
17. **Caldwell M., Andrews J. T. A., Tanay T., Griffin L. D.** AI-enabled future crime. *Crime Science*, 2020, vol. 9, no. 1, pp. 1–13. doi:10.1186/s40163-020-00123-8
18. **Schneier B.** Attack trees. *Dr. Dobb's Journal*, 1999, vol. 24, no. 12, pp. 21–29.
19. **Hutchins E. M., Cloppert M. J., Amin R. M**. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*. Academic Publishing International Limited, 2011. Vol. 1. Pp. 80–106.
20. *MITRE. 2017. ATT & CK Matrix for Enterprise*. https://attack.mitre.org/matrices/enterprise/ (дата обращения: 21.05.2022).
21. **Bu Z.** *Zero-day attacks are not the same as zero-day vulnerabilities*. www.fireeye.com/blog/executive-perspective/2014/04/zero-day-attacks-are-not-the-same-as-zero-day-vulnerabilities.html (дата обращения: 12.01.2022).
22. **Di Pinto A., Dragoni Y., Carcano A.** *TRITON: The first ICS cyber attack on safety instrument systems*. Black Hat, USA, 2018. 26 p.
23. **Găitan V. G., Zagan I.** Experimental implementation and performance evaluation of an IoT access gateway for the modbus extension. *Sensors*, 2021, vol. 21, 246, iss. 1, pp. 1–24. doi:10.3390/s21010246
24. **Radoglou-Grammatikis P., Sarigiannidis P., Efstathopoulos G., Karipidis P., Sarigiannidis A.** DIDEROT: an intrusion detection and prevention system for DNP3-based SCADA systems. *Proc. of the 15th Intern. Conf. on Availability, Reliability and Security*, 2020, pp. 1–8. doi:10.1145/3407023.3409314
25. **Ustun T. S., Hussain S. M. S.** IEC 61850 Modeling of UPFC and XMPP communication for power management in microgrids. *IEEE Access*, 2020, vol. 8, pp. 141696–141704. doi:10.1109/ACCESS.2020.3013264
26. **Saranya T., Sridevi S., Deisy C., Tran Duc Chung, Ahamed Khan M. K. A.** Performance analysis of machine learning algorithms in intrusion detection system: A review. *Procedia Computer Science*, 2020, vol. 171, pp. 1251–1260. doi:10.1016/j.procs.2020.04.133
27. **Kumar I., Mohd N., Bhatt C., Sharma S. K.** Development of IDS using supervised machine learning. *Soft computing: Theories and applications*. M. Pant et al (eds.). Springer, 2020. Pp. 565–577. doi:10.1007/978-981-15-4032-5_52
28. **Sudar K. M., Deepalakshmi P.** Comparative study on IDS using machine learning approaches for software defined networks. *Intern. Journal of Intelligent Enterprise*, 2020, vol. 7, iss. 1–3, pp. 15–27. doi:10.1504/IJIE.2020.104642
29. **Leichtnam L., Totel E., Prigent N., Me L.** Sec-2graph: Network attack detection based on novelty detection on graph structured data. *Intern. Conf. on Detection of Intrusions and Malware, and Vulnerability Assessment*, Lisboa, Portugal, 2020, pp. 238–258. doi:10.1007/978-3-030-52683-2_12
30. **Ковтун Л. И., Крюков О. В., Саушев А. В., Антоненко С. И.** Аналитико-статистический метод оценки состояния и прогнозирования рисков сложных технических систем. *Надежность и качество: тр. Междунар. симп.*, Пенза, 25–31 мая 2020 г., 2020, т. 1, с. 264–269.
31. **Murray J. F., Hughes G. F., Kreutz-Delgado F.** Machine learning methods for predicting failures in hard drives: A multiple-instance application. *Journal of Machine Learning Research*, 2005, vol. 6, pp. 783–816. doi:10.5555/1046920.1088699
32. **Gnidko K. O., Dudkin A. S., Ivanov O. S., Lokhvitsky V. A., Pilkevich S. V., Sabirov T. R.** The Unconscious response of multimedia content consumers to emociogenic visual symbols: Experimental study and oculometry dataset. *Solid State Technology*, 2020, vol. 63, no. 6, pp. 4549–4558.
33. **Zhang S., Bahrampour S., Ramakrishnan N., Schott L., Shah M.** Deep learning on symbolic representations for large-scale heterogeneous time-series event prediction. *42nd Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 5–9, 2017, IEEE, 2017, pp. 5970–5974. doi:10.1109/ICASSP.2017.7953302
34. **Khreich W., Khosravifar B., Hamou-Lhadj A., Talhib C.** An anomaly detection system based on variable N-gram features and one-class SVM. *Information and Software Technology*, 2017, vol. 91, pp. 186–197. doi:10.1016/j.infsof.2017.07.009
35. **Timoneda J. C.** Estimating group fixed effects in panel data with a binary dependent variable: how the LPM outperforms logistic regression in rare events data. *Social Science Research*, 2021, vol. 93, Article 102486. doi:10.1016/j.ssresearch.2020.102486
36. **Cafiso S., Pappalardo G.** Safety effectiveness and performance of lane support systems for driving assistance and automation — Experimental test and logistic regression for rare events. *Accident Analysis & Prevention*, 2020, vol. 148, Article 105791. doi:10.1016/j.aap.2020.105791
37. **Xu J., Denman S., Fookes C., Sridharan S.** Detecting rare events using Kullback — Leibler divergence. *40th Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, April 19–24, 2015, IEEE, 2015, pp. 1305–1309. doi:10.1109/ICASSP.2015.7178181
38. **Cheon S. P., Kim S., Lee S-Y., Lee C-B.** Bayesian networks based rare event prediction with sensor data. *Knowledge-Based Systems*, 2009, vol. 22, iss. 5, pp. 336–343. doi:10.1016/j.knosys.2009.02.004

39. **Bourinet J. M.** Rare-event probability estimation with adaptive support vector regression surrogates. *Reliability Engineering & System Safety*, 2016, vol. 150, pp. 210–221. doi:10.1016/j.ress.2016.01.023
40. **Gong D., Liu L., Le V., Saha B., Mansour M. R., Venkatesh S., Hengel A.** Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. *The IEEE/CVF Intern. Conf. on Computer Vision*, 2019, pp. 1705–1714. doi:10.48550/arXiv.1904.02639
41. **Han L. I.** Using a dynamic K-means algorithm to detect anomaly activities. *Seventh Intern. Conf. on Computational Intelligence and Security*, Dec. 3–4, 2011, Sanya, China, IEEE, 2011, pp. 1049–1052. doi:10.1109/CIS.2011.233
42. **Pradhan M., Pradhan S. K., Sahu S. K.** Anomaly detection using artificial neural network. *International Journal of Engineering Sciences & Emerging Technologies*, 2012, vol. 2, no. 1, pp. 29–36. doi:10.23919/FRUCT.2017.8071288
43. **Naseer S., Saleem Y., Khalid S., Bashir M. K., Han J., Iqbal M. M., Han K.** Enhanced network anomaly detection based on deep neural networks. *IEEE Access*, 2018, vol. 6, pp. 48231–48246. doi:10.1109/ACCESS.2018.2863036
44. **Lindemann B., Maschler B., Sahlab N., Weyrich M.** A survey on anomaly detection for technical systems using LSTM networks. *Computers in Industry*, 2021, vol. 131, Article 103498. doi:10.1016/j.compind.2021.103498
45. **Aguayo L., Barreto G. A.** Novelty detection in time series using self-organizing neural networks: A comprehensive evaluation. *Neural Processing Letters*, 2018, vol. 47, pp. 717–744. doi:10.1007/s11063-017-9679-2
46. **Jodelka O., Anagnostopoulos C., Kolomvatsos K.** Adaptive novelty detection over contextual data streams at the edge using one-class classification. *12th Intern. Conf. on Information and Communication Systems (ICICS)*, May 24–26, 2021, Valencia, Spain, IEEE, 2021, pp. 213–219. doi:10.1109/ICICS52457.2021.9464585
47. **Kulkarni P. G., Praneet S. Y., Bongole R., Das B.** Deep detection of anomalies in static attributed graph. *Intern. Conf. on Machine Learning, Image Processing, Network Security and Data Sciences*, Singapore, Springer, 2020, pp. 627–640. doi:10.1007/978-981-15-6318-8_50
48. **Li L., Yan J., Wang H., Jin Y.** Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, vol. 32, iss. 3, pp. 1177–1191. doi:10.48550/arXiv.2102.01331
49. **Chalapathy R., Chawla S.** Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407, 2019.
50. **Yuan Y., Kaklamanos G., Hogrefe D**. A novel semi-supervised adaboost technique for network anomaly detection. *Proc. of 19th ACM Intern. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2016, pp. 111–114. doi:10.1145/2988287.2989177
51. **Alippi C., Roveri M., Trovo F.** A self-building and cluster-based cognitive fault diagnosis system for sensor networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, vol. 25, iss. 6, pp. 1021–1032. doi:10.1109/TNNLS.2014.2303651
52. *База угроз безопасности информации ФСТЭК*. https://bdu.fstec.ru/threat (дата обращения: 21.05.2022).
53. **Goh J., Adepu S., Junejo K. N., Mathur A.** A dataset to support research in the design of secure water treatment systems. *Intern. Conf. "Critical Information Infrastructures Security"*, Springer, Cham, 2016, pp. 88–99. doi:10.1007/978-3-319-71368-7_8

UDC 004.89
doi:10.31799/1684-8853-2022-4-29-43

A study of ways and solutions to increase the efficiency of detecting computer attacks on the objects of critical information infrastructure

V. N. Kuzmin[a], Dr. Sc., Mil., Professor, orcid.org/0000-0002-6411-4336, vka@mil.ru
A. B. Menisov[a], PhD, Tech., Researcher, orcid.org/0000-0002-9955-2694
[a]A. F. Mozhaiskiy Military Space Academy, 13, Zhdanovskaia Emb., 197198, Saint-Petersburg, Russian Federation

Introduction: In the era of information technology almost all organizations face a wide range of automated and rapidly spreading cyber threats. This is due not only to the growing complexity, diversity and scale of digitalization, but also to the enlargement of cyber threats and the area of their possible implementation. **Purpose:** To compare possible ways of improving the effectiveness of attack detection for the objects of critical information infrastructure (CII): to detect a rare event, anomaly or novelty in the functions of the objects of CII. **Results:** The principle of operation of the proposed (effective) approach to cyberattack detection is to identify and separate anomalies from normal functioning of objects with the use of the concept of dynamic change of labels for a variable class over time. Dynamic novelty detection is compared to other approaches in terms of F1-score. For SWaT data, which is a layout of a critical information infrastructure object as an automated control system, it was determined that attack detection improved by up to 7% using the proposed approach. **Practical relevance:** The results of the research have shown a reduction in the risk of conducting (developing) a computer attack on critical information infrastructure objects. A possible targeted application of the dynamic novelty detection approach is to optimize the

means of protecting information at critical information infrastructure facilities, as well as to integrate the proposed approach into the information security system as an intelligent detector.

Keywords — information security, artificial intelligence technologies, critical information infrastructure, neural networks.

For citation: Kuzmin V. N., Menisov A. B. A study of ways and solutions to increase the efficiency of detecting computer attacks on the objects of critical information infrastructure. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 29–43 (In Russian). doi:10.31799/1684-8853-2022-4-29-43

Financial support

The work was carried out within the framework of the grant of the President of the Russian Federation for state support of young Russian scientists — candidates of sciences MK-2485.2022.4.

References

1. Petrenko S. A. Cyber resilient platform for Internet of Things (IIoT/IoT) ed systems: survey of architecture patterns. *Voprosy kiberbezopasnosti*, 2021, no. 2, pp. 81–91. doi:10.21681/2311-3456-2021-2-81-91
2. Maynard P., McLaughlin K., Sezer S. Decomposition and sequential-AND analysis of known cyber-attacks on critical infrastructure control systems. *Journal of Cybersecurit*, 2020, vol. 6, iss. 1, pp. 1–20. doi:10.1093/cybsec/tyaa020
3. Langner R. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 2011, vol. 6, iss. 3, pp. 49–51. doi:10.1109/MSP.2011.67
4. Miller B., Rowe D. A survey SCADA of and critical infrastructure incidents. *Proc. of the 1st Annual Conference on Research in Information Technology*, New York, Oct. 2012, NY, United States, 2012, pp. 51–56. doi:10.1145/2380790.2380805
5. Assante M. J., Lee R. M. *The Industrial Control System Cyber Kill Chain*. SANS Institute InfoSec Reading Room, 2015. Vol. 1. 24 p.
6. Rrushi J., Farhangi H., Howey C., Carmichael K., Dabell J. A quantitative evaluation of the target selection of havex ics malware plugin. *Industrial Control System Security (ICSS) Workshop*, Los Angeles, California, USA, 2015, pp. 1–5.
7. *ICS-CERT A. Ongoing sophisticated malware campaign compromising ICS*. Available at: http://www.ics-cert.uscert.gov/alerts/ICS-ALERT-14-281-01B (accessed 12 January 2022).
8. Rrushi J. L. *Multi-range Decoy I/O defense of electrical substations against industrial control system malware*. In: *Resilience of Cyber-Physical Systems: From Risk Modelling to Threat Counteraction*. F. Flammini (eds). Springer, Cham., 2019. Pp. 151–175. doi:10.1007/978-3-319-95597-1_7
9. Lee R. M., Assante M. J., Conway T. German steel mill cyber attack. *Industrial Control Systems*, 2014, vol. 30, iss. 62, pp. 1–15.
10. Bencsáth B., Pék G., Buttyán L., Félegyházi M. *Duqu: A Stuxnet-like malware found in the wild*. CrySyS Lab Technical Report, 2011. Vol. 14. 60 p.
11. Zetter K. *A cyberattack has caused confirmed physical damage for the second time ever*. Available at: https://www.wired.com/2015/01/german-steel-mill-hack-destruction/ (accessed 12 January 2022).
12. Alladi T., Chamola V., Zeadally S. Industrial control systems: Cyberattack trends and countermeasures. *Computer Communications*, 2020, vol. 155, pp. 1–8. doi:10.1016/j.comcom.2020.03.007
13. Saenko I. B., Lauta O. S., Karpov M. A., Kribel A. M. Model of threats to information and telecommunication network resources as a key asset of critical infrastructure. *Electrosvyaz*, 2021, no. 1, pp. 36–44 (In Russian). doi:10.34832/ELSV.2021.14.1.004
14. Rybkina O. V. Building a model of information security threats based on the Lancaster mathematical model. *Trudy Vserossijskoj nauchno-prakticheskoj konferencii "Nauchno-tekhnicheskoe i ekonomicheskoe sotrudnichestvo stran ATR v XXI veke"* [Proc. of the All-Russian Scient. and Pract. Conf. "Scientific, technical and economic cooperation of the Asia-Pacific countries in the 21st century"]. Khabarovsk, 2021, vol. 1, pp. 252–257 (In Russian).
15. Brazhuk A. I. Technique of threat modeling of computer systems based on subject-oriented models. *Trudy Vserossijskoj nauchnoj konferencii "Informacionnye tekhnologii v modelirovanii i upravlenii: podhody, metody, resheniya"* [Proc. of the All-Russian Scient. Conf. "Information technologies in modeling and management: approaches, methods, solutions"]. Tolyatti, 2021, pp. 94–101 (In Russian).
16. Sukhanov I. D., Rybkina O. V. New approaches to modeling information security threats. *Trudy Vserossijskoj nauchno-prakticheskoj konferencii "Nauchno-tekhnicheskoe i ekonomicheskoe sotrudnichestvo stran ATR v XXI veke"* [Proc. of the All-Russian Scient. and Pract. Conf. "Scientific, technical and economic cooperation of the Asia-Pacific countries in the 21st century"]. Khabarovsk, 2021, vol. 1, pp. 277–282 (In Russian).
17. Caldwell M., Andrews J. T. A., Tanay T., Griffin L. D. AI-enabled future crime. *Crime Science*, 2020, vol. 9, no. 1, pp. 1–13. doi:10.1186/s40163-020-00123-8
18. Schneier B. Attack trees. *Dr. Dobb's Journal*, 1999, vol. 24, no. 12, pp. 21–29.
19. Hutchins E. M., Cloppert M. J., Amin R. M. *Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains*. In: *Leading Issues in Information Warfare & Security Research*. Academic Publishing International Limited, 2011. Vol. 1. Pp. 80–106.
20. *MITRE. 2017. ATT & CK Matrix for Enterprise*. Available at: https://attack.mitre.org/matrices/enterprise/ (accessed 21 May 2022).
21. Bu Z. *Zero-day attacks are not the same as zero-day vulnerabilities*. Available at: www.fireeye.com/blog/executive-perspective/2014/04/zero-day-attacks-are-not-the-same-as-zero-day-vulnerabilities.html (accessed 12 January 2022).
22. Di Pinto A., Dragoni Y., Carcano A. *TRITON: The first ICS cyber attack on safety instrument systems*. Black Hat, USA, 2018. 26 p.
23. Găitan V. G., Zagan I. Experimental implementation and performance evaluation of an IoT access gateway for the modbus extension. *Sensors*, 2021, vol. 21, 246, iss. 1, pp. 1–24. doi:10.3390/s21010246
24. Radoglou-Grammatikis P., Sarigiannidis P., Efstathopoulos G., Karipidis P., Sarigiannidis A. DIDEROT: an intrusion detection and prevention system for DNP3-based SCADA systems. *Proc. of the 15th Intern. Conf. on Availability, Reliability and Security*, 2020, pp. 1–8. doi:10.1145/3407023.3409314
25. Ustun T. S., Hussain S. M. S. IEC 61850 Modeling of UPFC and XMPP communication for power management in microgrids. *IEEE Access*, 2020, vol. 8, pp. 141696–141704. doi:10.1109/ACCESS.2020.3013264
26. Saranya T., Sridevi S., Deisy C., Tran Duc Chung, Ahamed Khan M. K. A. Performance analysis of machine learning algorithms in intrusion detection system: A review. *Procedia Computer Science*, 2020, vol. 171, pp. 1251–1260. doi:10.1016/j.procs.2020.04.133
27. Kumar I., Mohd N., Bhatt C., Sharma S. K. *Development of IDS using supervised machine learning*. In: *Soft computing: Theories and applications*. M. Pant et al (eds.). Springer, 2020. Pp. 565–577. doi:10.1007/978-981-15-4032-5_52
28. Sudar K. M., Deepalakshmi P. Comparative study on IDS using machine learning approaches for software defined networks. *International Journal of Intelligent Enterprise*, 2020, vol. 7, iss. 1–3, pp. 15–27. doi:10.1504/IJIE.2020.104642
29. Leichtnam L., Totel E., Prigent N., Me L. Sec2graph: Network attack detection based on novelty detection on graph structured data. *Intern. Conf. on Detection of Intrusions and Malware, and Vulnerability Assessment*, Lisboa, Portugal, 2020, pp. 238–258. doi:10.1007/978-3-030-52683-2_12
30. Kovtun L. I., Kryukov O. V., Saushev A. V., Antonenko R. P. I. Analytical and statistical method for assessing the state and predicting the risks of complex technical systems. *Trudy Mezhdunarodnogo simpoziuma "Nadezhnost i kachestvo"* [Proc. Int. Symp. "Reliability and quality"], Penza, 2020, vol. 1, pp. 264–269 (In Russian).

31. Murray J. F., Hughes G. F., Kreutz-Delgado F. Machine learning methods for predicting failures in hard drives: A multiple-instance application. *Journal of Machine Learning Research*, 2005, vol. 6, pp. 783–816. doi:10.5555/1046920.1088699
32. Gnidko K. O., Dudkin A. S., Ivanov O. S., Lokhvitsky V. A., Pilkevich S. V., Sabirov T. R. The unconscious response of multimedia content consumers to emociogenic visual symbols: Experimental study and oculometry dataset. *Solid State Technology*, 2020, vol. 63, no. 6, pp. 4549–4558.
33. Zhang S., Bahrampour S., Ramakrishnan N., Schott L., Shah M. Deep learning on symbolic representations for large-scale heterogeneous time-series event prediction. *42nd Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 5970–5974. doi:10.1109/ICASSP.2017.7953302
34. Khreich W., Khosravifar B., Hamou-Lhadj A., Talhib C. An anomaly detection system based on variable N-gram features and one-class SVM. *Information and Software Technology*, 2017, vol. 91, pp. 186–197. doi:10.1016/j.infsof.2017.07.009
35. Timoneda J. C. Estimating group fixed effects in panel data with a binary dependent variable: how the LPM outperforms logistic regression in rare events data. *Social Science Research*, 2021, vol. 93, Article 102486. doi:10.1016/j.ssresearch.2020.102486
36. Cafiso S., Pappalardo G. Safety effectiveness and performance of lane support systems for driving assistance and automation — Experimental test and logistic regression for rare events. *Accident Analysis & Prevention*, 2020, vol. 148, Article 105791. doi:10.1016/j.aap.2020.105791
37. Xu J., Denman S., Fookes C., Sridharan S. Detecting rare events using Kullback — Leibler divergence. *40th Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 1305–1309. doi:10.1109/ICASSP.2015.7178181
38. Cheon S. P., Kim S., Lee S-Y., Lee C-B. Bayesian networks based rare event prediction with sensor data. *Knowledge-Based Systems*, 2009, vol. 22, iss. 5, pp. 336–343. doi:10.1016/j.knosys.2009.02.004
39. Bourinet J. M. Rare-event probability estimation with adaptive support vector regression surrogates. *Reliability Engineering & System Safety*, 2016, vol. 150, pp. 210–221. doi:10.1016/j.ress.2016.01.023
40. Gong D., Liu L., Le V., Saha B., Mansour M. R., Venkatesh S., Hengel A. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. *The IEEE/CVF Intern. Conf. on Computer Vision*, 2019, pp. 1705–1714. doi:10.48550/arXiv.1904.02639
41. Han L. I. Using a dynamic K-means algorithm to detect anomaly activities. *Seventh Intern. Conf. on Computational Intelligence and Security*, Dec. 3–4, 2011, Sanya, China, IEEE, 2011, pp. 1049–1052. doi:10.1109/CIS.2011.233
42. Pradhan M., Pradhan S. K., Sahu S. K. Anomaly detection using artificial neural network. *International Journal of Engineering Sciences & Emerging Technologies*, 2012, vol. 2, no. 1, pp. 29–36. doi:10.23919/FRUCT.2017.8071288
43. Naseer S., Saleem Y., Khalid S., Bashir M. K., Han J., Iqbal M. M., Han K. Enhanced network anomaly detection based on deep neural networks. *IEEE Access*, 2018, vol. 6, pp. 48231–48246. doi:10.1109/ACCESS.2018.2863036
44. Lindemann B., Maschler B., Sahlab N., Weyrich M. A survey on anomaly detection for technical systems using LSTM networks. *Computers in Industry*, 2021, vol. 131, Article 103498. doi:10.1016/j.compind.2021.103498
45. Aguayo L., Barreto G. A. Novelty detection in time series using self-organizing neural networks: A comprehensive evaluation. *Neural Processing Letters*, 2018, vol. 47, pp. 717–744. doi:10.1007/s11063-017-9679-2
46. Jodelka O., Anagnostopoulos C., Kolomvatsos K. Adaptive novelty detection over contextual data streams at the edge using one-class classification. *12th Intern. Conf. on Information and Communication Systems (ICICS)*, 2021, pp. 213–219. doi:10.1109/ICICS52457.2021.9464585
47. Kulkarni P. G., Praneet S. Y., Bongole R., Das B. Deep detection of anomalies in static attributed graph. *Intern. Conf. on Machine Learning, Image Processing, Network Security and Data Sciences*, Singapore, Springer, 2020, pp. 627–640. doi:10.1007/978-981-15-6318-8_50
48. Li L., Yan J., Wang H., Jin Y. Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, vol. 32, iss. 3, pp. 1177–1191. doi:10.48550/arXiv.2102.01331
49. Chalapathy R., Chawla S. Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407, 2019.
50. Yuan Y., Kaklamanos G., Hogrefe D. A novel semi-supervised adaboost technique for network anomaly detection. *19th ACM Intern. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2016, pp. 111–114. doi:10.1145/2988287.2989177
51. Alippi C., Roveri M., Trovo F. A self-building and cluster-based cognitive fault diagnosis system for sensor networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, vol. 25, iss. 6, pp. 1021–1032. doi:10.1109/TNNLS.2014.2303651
52. *Baza ugroz bezopasnosti informacii FSTEK* [FSTEC information security threat database]. Available at: https://bdu.fstec.ru/threat (accessed 21 May 2022).
53. Goh J., Adepu S., Junejo K. N., Mathur A. A dataset to support research in the design of secure water treatment systems. *Intern. Conf. "Critical Information Infrastructures Security"*, Springer, Cham, 2016, pp. 88–99. doi:10.1007/978-3-319-71368-7_8

УДК 004.056.5
doi:10.31799/1684-8853-2022-4-44-57

Научные статьи
Articles

Применение алгоритмов машинного обучения для обнаружения вредоносных программ в операционной системе Windows с помощью PE-заголовка

Д. Ч. Ле*[а], *канд. техн. наук, преподаватель, orcid.org/0000-0003-3735-0314, letranduc@dut.udn.vn
М. Х. Фам*[б], *инженер, orcid.org/0000-0002-2250-9428
Ч. З. Динь*[в], *канд. техн. наук, преподаватель, orcid.org/0000-0002-9993-9792
Х. Ф. До*[г], *магистр, преподаватель, orcid.org/0000-0003-0645-0021
[а]*Университет Дананга — Университет науки и техники, факультет информационных технологий, 54 Nguyen Luong Bang, Дананг, 550000, Вьетнам*
[б]*Отдел исследований и разработок, Viettel Business Solutions Корпорация, Дананг, Вьетнам*
[в]*Технологический институт почты и телекоммуникаций, 122 Hoang Quoc Viet, Ханой, Вьетнам*
[г]*Университет архитектуры Дананга, 566 Nui Thanh, Дананг, Вьетнам*

Введение: *быстрый рост числа вредоносных программ и их злонамеренное использование приводят к значительным финансовым потерям. Исследователи заинтересованы в применении методов машинного обучения для решения задачи обнаружения вредоносных программ. Однако в силу разнообразия алгоритмов машинного обучения каждый имеет свой подход в определенной ситуации.* ***Цель:*** *применить методы машинного обучения для обнаружения вредоносных программ в операционной системе Windows с использованием компонентов переносимого исполняемого (Portable Executable) заголовка; сравнить шесть алгоритмов машинного обучения по нескольким критериям.* ***Результаты:*** *сравнение различных алгоритмов, таких как случайный лес (Random Forest), дерево принятия решений (Decision Tree), наивный байесовский алгоритм (Naive Bayes), метод опорных векторов (Support Vector Machine), многослойный перцептрон (Multilayer Perceptron), метод k-ближайших соседей (k-Nearest Neighbors) с большим набором данных, показало, что применение алгоритмов случайный лес, дерево принятия решений, метод k-ближайших соседей и многослойный перцептрон позволяет довести вероятность обнаружения вредоносных программ до высокой точности (> 98 %). Особенно алгоритм случайный лес очень хорошо подходит для применения в средствах обнаружения вредоносного программного обеспечения на операционной системе Windows. Наивный байесовский алгоритм также имеет высокий показатель точности (> 96 %) и быстрое время обработки. Поэтому мы можем рассматривать возможность использовать наивный байесовский алгоритм в качестве альтернативного.*

Ключевые слова *— вредоносная программа, алгоритм машинного обучения, PE-заголовок, Windows.*

Для цитирования: Ле Д. Ч., Фам М. Х., Динь Ч. З., До Х. Ф. Применение алгоритмов машинного обучения для обнаружения вредоносных программ в операционной системе Windows с помощью PE-заголовка. *Информационно-управляющие системы*, 2022, № 4, с. 44–57. doi:10.31799/1684-8853-2022-4-44-57

For citation: Le D. T., Pham M. H., Dinh T. D., Do H. P. Applying machine learning algorithms for PE-header-based malware detection on the Windows operating system. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 44–57 (In Russian). doi:10.31799/1684-8853-2022-4-44-57

Введение

В настоящее время с быстрым развитием Интернета компьютерные приложения и системное программное обеспечение (ПО) постоянно меняются и развиваются. Это сопровождается появлением различных типов вредоносных программ [1]. Сегодня вредоносные программы становятся все более разнообразными, сложными и опасными [2].

Вредоносные программы являются одной из прямых угроз информационной безопасности. Эти вредоносные файлы распространяются незаметно и быстро различными передовыми методами, такими как механизмы инъекций, антивиртуальная машина, антиотладка, упаковщик, шифрование и механизмы настойчивости [3, 4].

Вредоносные программы идентифицируются и классифицируются по их вредоносной цели или поведению [5]. Например, различают такие типы вредоносных программ, как вирусы, трояны, программы-вымогатели, бэкдоры, рекламные программные обеспечения и т. д.

В настоящее время существует много инструментов для анализа вредоносных программ: *PEStudio*, *CFF Explorer*, *DetecItEasy*, *PPEE*, *IDA*, *Ghidra*, *OllyDbg*, *x64dbg*, *radare2*, *WinDbg*, *gdb* [6]. Их использует человек, осуществляющий анализ вредоносного ПО, но работать автономно, выявляя вредоносные программы, они не способны. Кроме того, каждый инструмент имеет свои преимущества и недостатки, а результат в основном зависит от опыта человека, проводящего анализ. Можно использовать методы статического

анализа вредоносных программ (*static malware analysis*), чтобы найти некоторые важные функции и спрогнозировать поведение подозрительных файлов, но это требует много времени. Также можно использовать динамический анализ в режиме реального времени, когда подозрительный файл выполняется в виртуальной среде, а по признакам при запуске файла в операционной системе (ОС) можно узнать, является ли он вредоносным или нет. Однако у динамического анализа много недостатков, например: для каждого типа вредоносного ПО требуется соответствующая среда анализа; легко обнаруживается функцией антивиртуальной машины вредоносного ПО и т. д.

При большом количестве вредоносных программ вышеперечисленные методы сложны для реализации. Для того чтобы сократить время анализа, необходимо иметь автоматическую систему классификации вредоносных программ. В этой статье мы предлагаем использовать алгоритмы машинного обучения для обнаружения вредоносных программ в ОС Windows. В частности, мы фокусируемся на формате PE-заголовка с существенными функциями/признаками, которые помогают обнаруживать много вредоносных ПО в Windows. Многие исследователи заинтересованы в применении методов машинного обучения для решения различных проблем в анализе вредоносных программ [7–9]. Однако из-за разнообразия каждый алгоритм машинного обучения имеет свой подход в определенной ситуации. Основное внимание в этой статье сосредоточено на применении машинного обучения для обнаружения вредоносных программ на Windows. Ко всему прочему, в статье представлено сравнение шести различных алгоритмов машинного обучения на основе нескольких критериев для расширения выбора и поиска оптимального решения. Этими алгоритмами являются случайный лес (*Random Forest*, *RF*), дерево принятия решений (*Decision Tree*, *DT*), наивный байесовский алгоритм (*Naive Bayes*, *NB*), метод опорных векторов (*Support Vector Machine*, *SVM*), многослойный перцептрон (*Multilayer Perceptron*, *MLP*), метод *k*-ближайших соседей (*k-Nearest Neighbors*, *k-NN*) [10], все они поддерживаются библиотеками Python Scikit-Learn [11]. Следует отметить, что многие категории вирусного ПО, нацеленного на ОС Windows, не имеют PE-заголовка. Однако рассмотрение этих типов выходит за рамки данной статьи.

Обзор литературы

В работе [12] авторы представили фреймворк для обнаружения вредоносных программ, направленный на получение как можно меньше ложноположительного результата, используя простую многоступенчатую комбинацию (каскад) различных версий алгоритма перцептрона. Перцептрон — это бинарный классификатор, который определяет принадлежность входных данных к определенному классу на основе набора весов с векторами признаков. Следует отметить, что алгоритмы, которые привели к лучшей точности, также дали наибольшее количество ложноположительного результата. Предлагаемая схема основана на статическом анализе, который может занять много времени.

В статье [13] Х. Ратхор и др. предложили глубокое обучение на основе алгоритма случайного леса с опкодом в качестве вектора признаков для обнаружения вредоносных программ. Они разобрали исполняемый файл с помощью *objdump* и извлекли опкод из файла. В качестве данных признаков был составлен и использован главный список опкодов, состоящий из 1600 уникальных опкодов. Однако в этой статье для обнаружения вредоносных программ были созданы только алгоритм случайного леса и различные слои глубокой нейронной сети.

Авторы в статье [14] использовали динамические характеристики и шесть алгоритмов машинного обучения, чтобы обнаружить варианты (на основе уязвимости нулевого дня) версий программ-вымогателей среди вредоносных и нормальных приложений. Экспериментальные результаты показали, что предлагаемый метод может обнаружить программы-вымогатели среди вредоносных программ и доброкачественных файлов.

Аша Джерлин и Маримутху Каруппиа [15] разработали эффективную систему для обнаружения вредоносных программ с использованием программируемых интерфейсов приложений (APIs) и классификации их типа как червей, вирусов и троянов или доброкачественных файлов. Предварительно набор входных данных обработан путем нормализации данных, затем его верхние и нижние границы оценивались во время выделения признаков. Результаты экспериментов показали, что предлагаемые методы хорошо работают при проверке с использованием таких критериев, как чувствительность алгоритма классификации (*True Positive Rate*, *TPR*), специфичность алгоритма классификации (*False Positive Rate*, *FPR*), точность (*precision*), полнота (*recall*), F1-мера (F1-*score*) и точность (*accuracy*).

Чандрасекар Рави и Р. Манохаран [16] также предложили систему обнаружения вредоносных программ, которая использует набор базовых функций интерфейсов программирования приложений ОС Windows API. Для моделирования вызовов API они использовали цепочку Маркова 3-го порядка. В данной статье используется классификация на основе ассоциативного майнинга, поскольку она дает более высокую точность

(*accuracy*) обнаружения. Вредоносные файлы в основном состояли из бэкдоров, червей и троянских коней, собранных из VXHeavens.

В работе [17] авторы попытались построить систему обнаружения вредоносных программ на основе их поведения. Модель объединила метод опорных векторов и метод главных компонент.

Джисин Сюй и др. предложили решение для обнаружения вредоносных программ с аппаратной поддержкой, которое использовало методы машинного обучения для мониторинга виртуальной памяти на наличие вредоносных доступов, вызванных вредоносными программами [18]. Новые аспекты фреймворка включают методы сбора и обобщения паттернов доступа к памяти для каждой функции/системного вызова и двухуровневую архитектуру классификации. Однако авторы обучали по одной модели для каждого приложения, что может быть довольно дорого.

Одним из наиболее часто используемых решений обнаружения вредоносных программ, особенно в антивирусном ПО, является обнаружение на основе сигнатур (*signature*). Недостаток этого решения заключается в том, что оно не может обнаружить новые вредоносные программы. Современные вредоносные программы имеют несколько полиморфных слоев, чтобы избежать обнаружения или автоматически обновить себя до новой версии. В работе [19] авторы предложили использовать n-граммы в качестве сигнатур файлов для обнаружения неизвестных вредоносных программ при сохранении низкого ложноположительного соотношения. Результаты показали, что n-граммовые сигнатуры обеспечивают эффективный способ обнаружения неизвестных вредоносных программ.

В отличие от обнаружения на основе сигнатур, эвристическое сканирование использует правила и (или) алгоритмы для поиска команд, которые могут указывать на намерение и подозрительные знаки. При использовании этого метода можно обнаружить вредоносные программы без наличия сигнатур. Большинство антивирусных программ используют эвристические методы.

В нашей работе мы предлагаем метод на основе выделенных признаков из PE-заголовка. Поля заголовков подозрительных файлов выделяются и сравниваются с полями заголовков доброкачественных файлов. Мы также попытаемся обобщить PE-признаки, которые оказывают значительное влияние на обнаружение вредоносных программ в Windows.

Следует отметить, что использование информации, извлеченной из PE-заголовка, для обнаружении вредоносных программ упоминалось некоторыми исследователями ранее. Например, Цзиньронг Бай и др. [20] фокусировались на библиотеках DLL (*Dynamic Link Library*) и API-функциях, извлеченных из анализа таблицы импорта (*Import Table*) в PE-заголовке. Однако библиотеки DLL и API-функции часто меняются в зависимости от характеристик каждого типа вредоносного ПО. Поэтому необходимо проводить более тщательный и избирательный анализ важной информации PE-заголовка в процессе изучения вредоносных программ.

Другой проект [21] сосредоточен на анализе и использовании информации PE-заголовка в комбинации с машинным обучением в обнаружении вредоносных программ. Однако авторы не указали, как PE-заголовок влияет на точность и эффективность моделей машинного обучения. В статье [22] авторы предложили использовать как исходное, так и выделяемое значение из PE-заголовка для создания входных данных для модели классификации, при этом они не указывают степень влияния выделяемых признаков на критерии, используемые для оценки моделей.

Основная цель нашего исследования заключается в применении алгоритмов машинного обучения случайный лес, дерево принятия решений, наивный байесовский алгоритм, метод опорных векторов, многослойный перцептрон, метод k-ближайших соседей на основе признаков, извлеченных из PE-заголовка, для обнаружения вредоносных программ на ОС Windows. Также мы проводим сравнение различных алгоритмов на основе метрик для оценки модели обнаружения вредоносных программ, включая точность, время обработки и эффективность. Обычно для моделей машинного обучения, используемых в задачах классификации, точность (*accuracy*) является основным критерием для оценки. Между тем время обработки — это время для модели, чтобы классифицировать файл ввода как вредоносную программу, и оно показывает способность модели к масштабированию при больших наборах данных. Кроме того, чтобы более объективно оценить алгоритмы и иметь возможность выбрать правильную модель, мы полагаемся на информацию из матрицы ошибок, точности, полноты, F1-меры и кривой соотношений правильного и ложного обнаружения (ROC-кривой, *Receiver Operating Characteristic curve*), показывающей производительность и эффективность классификационной модели алгоритмов.

Вредоносные программы в ОС Windows и формат PE-заголовка

В настоящее время вредоносные программы появляются во всех ОС, и все ОС восприимчивы к вредоносным программам. Относительно легко найти вредоносные программы для каждой

ОС, от Android, iOS до macOS, Linux и Windows. Среди этих ОС Windows является самой популярной операционной системой в мире, и, следовательно, она стала объектом большинства вредоносных атак. По данным отчета о безопасности AV Test за 2019/2020 г. (https://www.av-test.org/fileadmin/pdf/security_report/AV-TEST_Security_Report_2019-2020.pdf), в 2019 г. было разработано 114 миллионов новых вредоносных программ, а 78,64 % всех атак было распространено на ОС Windows.

В большинстве предыдущих исследований, связанных с машинным обучением, авторы в основном извлекали признаки с помощью статического или динамического анализа из образцов вредоносных программ. При статическом анализе признаки извлекаются из последовательности байтов или значений опкода. В то же время при динамическом анализе признаки извлекаются путем запуска или эмуляции кода. Эти исследования в основном требуют выполнения всей программы и меньшего внимания к PE-заголовкам — одному из ключевых компонентов вредоносных программ, работающих в Windows-среде. Более того, запуск тысяч вредоносных программ для извлечения признаков, необходимых для процесса обучения, очень сложно реализовать.

Формат Portable Executable (PE) — это формат для выполнимых файлов, объектного кода, библиотек DLL и других файлов, используемых в 32-разрядных и 64-разрядных версиях ОС Windows. PE получен из спецификации Common Object File Format (COFF), которая также используется большинством выполнимых файлов Unix. Типы файлов, которые используют PE-формат, включают: .exe, .dll, .acm, .ax, .cpl, .drv, .efi, .mui, .ocx, .scr, .sys, .tsp.

PE-заголовок содержит информацию, которая считывается загрузчиком Windows при выполнении файла. После этого содержимое файла будет загружено из файла в память. С помощью PE-заголовка программа информирует ОС о своих требованиях к выполнению, поскольку он указывает, где исполняемый файл должен быть загружен в память. Он также указывает список библиотек/функций, на которые опирается приложение; адрес, с которого начинается выполнение, и двоичные ресурсы. Проверка PE-заголовка дает огромное количество информации о подозрительном файле и его функциях. Поэтому PE-заголовок имеет большое значение при обнаружении и анализе вредоносных программ. Основная структура PE-заголовка включает в себя:

— *DOS MZ Header*: проверяет, является ли файл действительным PE-файлом или нет (основные компоненты*:* e_magic, e_ifanew);

— *DOS Stub*: отображает предупреждение, если файл не может быть запущен в Windows;

— *PE File Header*: содержит информацию о файле (основные компоненты: signature, file header, optional header...);

— *Section Table*: содержит информацию о разделах, присутствующих в PE-файлах (основные компоненты: Name1, VirtualSize, SizeOfRawData, PointerToRawData, Characteristics);

— *Sections*: размещает основное содержимое файла, включая код, данные, ресурсы и другие исполняемые файлы (основные компоненты: .TEXT, .RDATA, .DATA, .RSRC).

В PE-файлах есть много признаков, но большинство из них не помогают отличить вредоносное ПО от доброкачественного. На основе наших эмпирических исследований и углубленного анализа признаков PE-заголовка мы извлекли 54 признака, отличающих доброкачественное ПО от вредоносных. Например: Machine, SizeOfOptionalHeader, Characteristics, MajorLinkerVersion, MinorLinkerVersion, SizeOfCode, AddressOfEntryPoint, ImageBase, Subsystem, DllCharacteristics и т. д. Полное описание этих признаков представлено в (https://docs.microsoft.com/ru-ru/windows/win32/debug/pe-format).

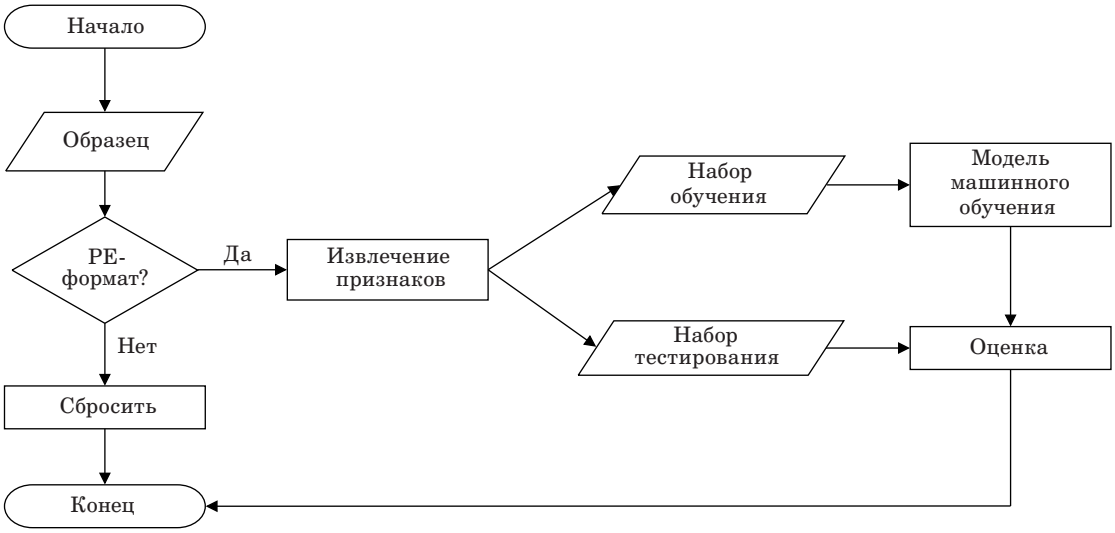
Экспериментальная часть

Рассмотрим этапы проведения экспериментов по оценке алгоритмов машинного обучения. На рис. 1 представлен алгоритм реализации.

Эксперимент проводился на Windows 10, 64-битный компьютер с процессорами Core AMD A8-4555M (1,6 ГГц, 4 ядра, 4 процессора), 8 ГБ оперативной памяти. Благодаря статическому анализу вычислительные затраты очень низки, поэтому все эксперименты проводятся на типичной системе конечных хостов.

Подготовка набора данных

Во-первых, необходимо получить набор данных для обучения алгоритмов. Для создания набора данных мы используем исполняемые файлы, зараженные вредоносными программами, предоставленные в Virusshare (https://virusshare.com/). Чтобы получить доступ к общей базе данных вредоносных программ, необходимо запросить учетную запись по электронной почте по адресу admin@virusshare.com. Набор данных вредоносных программ от Virusshare обычно не имеет расширения. Они будут переформатированы в исполняемые файлы автоматически с помощью команды *Windows Cmd*. В результате был использован набор данных, состоящий из 4761 файла, из которых 3816 были заражены



■ ***Рис. 1.*** Алгоритм реализации экспериментальной части работы
■ ***Fig. 1.*** Algorithm for the implementation of the experimental part

вредоносными программами, а 945 файлов были незараженными, доброкачественными. Следует отметить, что наш набор данных вредоносных программ имеет вредоносное ПО практически только для архитектуры x86-64. Результаты, полученные в эмпирическом процессе, могут отличаться для набора данных, содержащего вредоносное ПО для архитектуры x64.

Была использована *Python* — программа для извлечения признаков PE-заголовка. В этом эксперименте вредоносные программы помечаются как 1 (положительный ярлык) и 0 (отрицательный ярлык) для обычных файлов. После извлечения и маркировки всех файлов в допустимом формате получен набор данных для обучения и тестирования в моделях машинного обучения.

Предварительная переработка данных

Набор данных после сбора из PE-файлов был разделен на обучающий и тестовый наборы в соотношении 70 и 30 %. Набор обучения и набор тестирования были стандартизированы.

Нормализация — это хороший метод, который можно использовать, когда распределение данных неизвестно или оно не является гауссовым (колоколообразная кривая). Нормализация полезна, когда данные имеют различный масштаб, а используемый алгоритм не делает предположений о распределении данных, например *k-NN* и искусственные нейронные сети.

При нормализации данных среднее значение каждого признака равно нулю, а дисперсия равна 1. Для того чтобы формализовать данные, мы должны определить среднее и стандартное отклонение для распределения каждого признака. Затем вычесть из значения каждого признака среднее значение, далее разделить его на стандартное отклонение признака следующим образом:

$$x' = \frac{x - average(x)}{std(x)}, \qquad (1)$$

где x — исходный вектор признаков; $average(x)$ — среднее значение данного вектора признаков; $std(x)$ — стандартное отклонение.

Обучение модели

Нормализованные наборы данных будут обучаться в моделях машинного обучения. Прогнозируемые результаты моделей машинного обучения существенно зависят от гиперпараметров алгоритмов. Следовательно, очень важно выбрать экспериментальные значения для гиперпараметров в соответствии с каждой моделью, чтобы получить наилучшие результаты. Изначально модель обучается по конкретному алгоритму без внесения корректировки гиперпараметров (т. е. с использованием набора гиперпараметров по умолчанию в *scikit-learn*). После этого получены результаты с низкой точностью. По этим результатам было отмечено, что образцы всегда классифицируются в один класс. Следовательно, необходимо регулировать его параметры. Здесь использована *GridSearch* (в *scikit-learn*), чтобы регулировать гиперпараметры для этой модели. Далее эксперименты повторялись, и после настройки гиперпараметров получены лучшие результаты. Тот же процесс выполнен для других алгоритмов, и составлена таблица значений гиперпараметров (табл. 1).

■ ***Таблица 1.*** Гиперпараметры для каждой модели машинного обучения

■ ***Table 1.*** Hyperparameters for each machine learning model

Модель	Гиперпараметр	Описание
Дерево принятия решений (*DT*)	*criterion*	*Gini*
	max_depth	*3*
	min_samples_split	*2*
	min_samples_leaf	*1*
Случайный лес (*RF*)	*n_estimators*	*50*
	criterion	*Entropy*
	min_samples_split	*2*
	min_samples_leaf	*1*
Наивный байесовский алгоритм (*NB*)	*var_smoothing*	*1e-9*
Метод *k*-ближайших соседей (*k-NN*)	*n_neighbors*	*3*
	weights	*Uniform*
	leaf _size	*30*
	p	*2*
	metric	*Minkowski*
	algorithm	*Auto*
Метод опорных векторов (*SVM*)	*Kernel*	*Linear*
	Gamma	*scale*
	C	*1.0*
	Tol	*1e-3*
	max_iter	*-1*
Многослойный перцептрон (*MLP*)	*solver*	*Lbfgs*
	alpha	*1e-5*
	hidden_layer_sizes	*(5, 2)*

Оценка модели

После обучения моделей набор тестирования будет использоваться для проверки прогнозируемых результатов. С набором данных, состоящим только из небольшого количества образцов, мы можем проверить прогнозируемые результаты визуально. Однако, поскольку текущие наборы данных содержат значительное количество образцов, для обобщения результатов и оценки эффективности моделей классификации необходима матрица ошибок (*confusion matrix*). В данном эксперименте файлы с вредоносным ПО были помечены положительно (*positive*), а файлам без вредоносного ПО присваивались отрицательные метки (*negative*).

Матрица ошибок имеет следующую информацию.

TP (*True Positive* — правильно положительный): образцы с положительным ярлыком правильно классифицируются в положительный класс, т. е. общее количество вредоносных программ, которое прогнозируется как вредоносное ПО (истинное значение было «вредоносное», и модель прогнозировала «вредоносное»).

FN (*False Negative* — ложноотрицательный): положительно помеченные образцы ошибочно классифицируются в отрицательный класс, т. е. общее количество вредоносных программ, которые прогнозируются как доброкачественное ПО (истинное значение было «вредоносное», а модель прогнозировала «доброкачественное»).

FP (*False Positive* — ложноположительный): отрицательно помеченные образцы ошибочно классифицируются в положительный класс, т. е. общий доброкачественный файл, который предсказывается как вредоносное ПО (истинное значение было «доброкачественное», а модель прогнозировала «вредоносное»).

TN (*True Negative* — правильно отрицательный): отрицательно помеченные образцы правильно классифицируются в отрицательный класс, т. е. общий доброкачественный файл, который предсказывается как доброкачественное ПО (истинное значение было «доброкачественное», и модель прогнозировала «доброкачественное»).

Точность (или достоверность — *accuracy*) — это отношение правильно спрогнозированных наблюдений к общему количеству наблюдений. Точность модели показывает производительность модели и рассчитывается по следующей формуле:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \qquad (2)$$

Для оценки эффективности алгоритмов машинного обучения мы также можем использовать некоторые из следующих метрик: точность (*precision*), полноту (*recall*), ROC-кривую и F1-меру.

Точность (*precision* — *P*) является положительным прогностическим значением, которое может быть получено из матрицы ошибок. Она определяется как количество сделанных прогнозов, которые на самом деле являются правильными из всех прогнозов, основанных на положительном классе. Другими словами, это доля прогнозируемых вредоносных программ, которые на самом деле являются вредоносными программами. Точность является важной метрикой, поскольку неправильное предсказание вредоносного ПО опасно для реальных систем. Ее можно рассчитать по формуле

$$Precision = \frac{TP}{TP + FP}. \qquad (3)$$

Полнота (*recall — R*), также известна как чувствительность, или коэффициент *TP* (*TPR*), используется для определения соответствующих точек данных в процентах. Определяется как количество образцов положительного класса, которые были правильно спрогнозированы. Другими словами, это доля фактических вредоносных программ от прогнозируемых вредоносных программ. Значение полноты может быть рассчитано следующим образом:

$$Recall = \frac{TP}{TP + FN}. \quad (4)$$

ROC-кривая (*Receiver Operating Characteristic curve*) может быть использована для мультиклассовых классификаторов. Показатель *TP* (*TPR*) и показатель *FP* (*FPR*) классификатора используются для построения ROC-кривой. *FPR* также упоминается как ложные тревоги, определяя общее количество неправильно положительных предсказаний среди всех отрицательных образцов в наборе данных. Мы можем рассчитать *FPR* следующим образом:

$$FPR = \frac{FP}{TN + FP}. \quad (5)$$

F1-мера (F1-*score*) представляет собой комбинацию точности и полноты. Это гармоническое среднее между точностью и полнотой. Она принимает значение в диапазоне (0,1]. F1-мера рассчитывается следующим образом:

$$\text{F1-}score = \frac{2 * Recall * Precision}{Recall + Precision}. \quad (6)$$

Кроме того, более высокие значения точности, полноты и F1-меры указывают на эффективность классификации.

Полученные результаты и анализ моделей

В этом разделе сравним уровень влияния признаков на обучение моделей, потому что разные признаки играют разную роль в обучении. Экспериментально покажем, что существуют некоторые признаки, называемые лучшими, которые имеют гораздо большую степень влияния, чем другие.

Лучшие признаки, выбранные из 54 признаков PE-заголовка и отсортированные по степени влияния на процесс обнаружения вредоносного ПО, показаны в табл. 2

Мы видим, что признак *Machine* имеет самый значительный показатель воздействия — 23,21 %. Другие признаки хотя и не достигли самого высокого показателя, все же играют существенную роль в классификации вредоносных и доброкачественных файлов. В частности, на 9-й из 54 признаков приходится 79,63 % от показателя воздействия.

Использование лучших признаков из табл. 2 полностью оправдано, особенно при статическом анализе вредоносного ПО, поскольку они содержат важную информацию, необходимую вредоносному ПО для эффективного выполнения в ОС Windows.

Machine: это поле в заголовке файла *COFF*. Оно устанавливает тип архитектуры целевой машины, например Intel, AMD. Это поле частично определяет, какой процессор является целью текущего вредоносного ПО.

ImageBase: когда вредоносная программа выполняется, загрузчик Windows создает процесс для этой вредоносной программы. Загрузчик Windows копирует и загружает PE-файл и секции вредоносной программы с диска в виртуальную память процесса. Поле *ImageBase* указывает предпочтительный адрес, по которому исполняемый файл должен быть отображен в памяти. По умолчанию используется адрес 0x00400000 для 32-разрядных исполняемых файлов, для .dll он иной — 0x00100000.

MajorSubsystemVersion: обычно используется для проверки совместимости системы с текущим PE-файлом. Вредоносные программы часто используют это поле в сочетании с *MajorOperatingSystemVersion*, чтобы убедиться, что они выполняются на зараженной системе, поддерживающей максимум встроенных функций вредоносной программы.

■ ***Таблица 2.*** Лучшие признаки и соответствующий показатель воздействия

■ ***Table 2.*** Best features and the corresponding impact indicator

Ранг	Признак	Показатель воздействия, %
1	Machine	23,21
2	ImageBase	19,37
3	MajorSubsystemVersion	14,42
4	Characteristics	7,37
5	SizeOfOptionalHeader	3,74
6	MajorLinkerVersion	3,68
7	MajorOperatingSystemVersion	3,03
8	Subsystem	2,89
9	MinorLinkerVersion	1,92

Characteristics: указывает атрибуты объекта или файла изображения. Оно помогает определить, имеет ли вредоносная программа вид .exe или .dll, поскольку оба эти расширения используют формат PE-заголовка. По полю *Characteristics* с помощью флага *IMAGE_FILE_32BIT_MACHINE* мы также можем определить, нацелена ли эта вредоносная программа на 32-разрядную или 64-разрядную Windows.

SizeOfOptionalHeader: обеспечивает точный размер необязательного заголовка (*Optional Header*), который необходим для правильного анализа PE-файла. Кроме того, значение этого поля также влияет на сложность PE-файла, поскольку оно напрямую связано с размером других важных полей, таких как *ExportTable*, *ImportTable*, *ResourceTable*, *ImporAddressTable*.

MajorLinkerVersion и **MinorLinkerVersion**: указывают версию компоновщика, который используется для получения объектных файлов (сгенерированных компилятором или ассемблером) и объединения их в исполняемый файл. Обычно эти значения используются для сравнения с *Rich Headers* [23], чтобы найти упакованные файлы.

MajorOperatingSystemVersion: указывает минимальную версию ОС, необходимую для использования этого исполняемого файла.

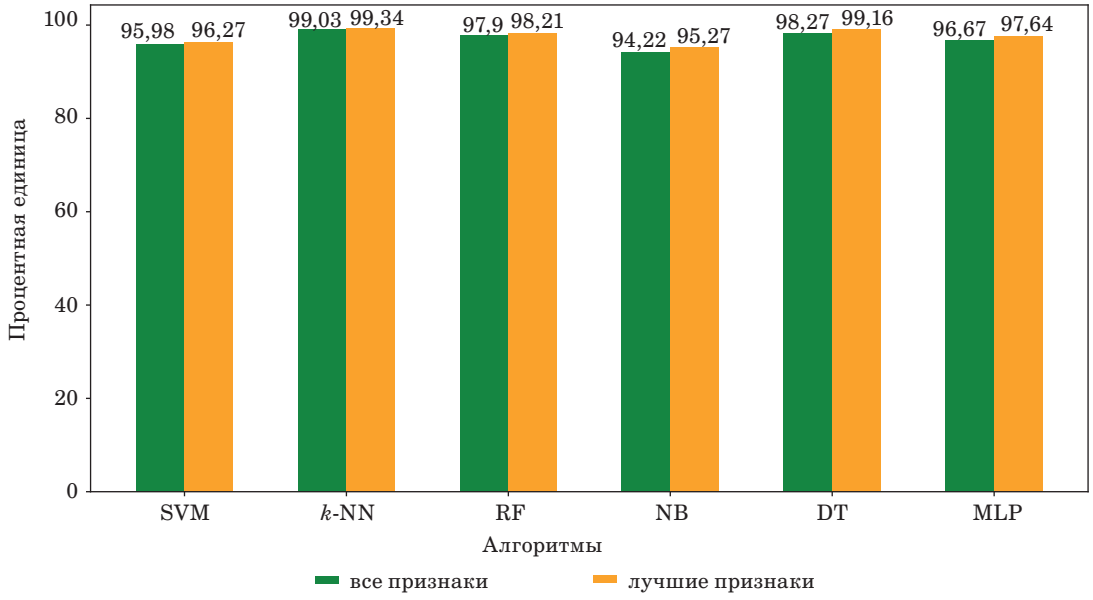
SubSystem: определяет, какая подсистема Windows (если она существует) требуется для запуска файла. Это поле указывает, что PE-файл будет запущен с помощью драйвера, графического или командного интерфейса пользователя.

Анализ

Мы пытались проверить, влияет ли использование лучших признаков на обнаружение вредоносных программ. На самом деле, нельзя сказать, что использование только лучших признаков статистически улучшило результаты. Однако благодаря многим экспериментам мы получили результаты, изложенные ниже. Следует отметить, что эти результаты могут отличаться в зависимости от набора данных, ОС и применимых алгоритмов.

Анализ на основе точности (accuracy)

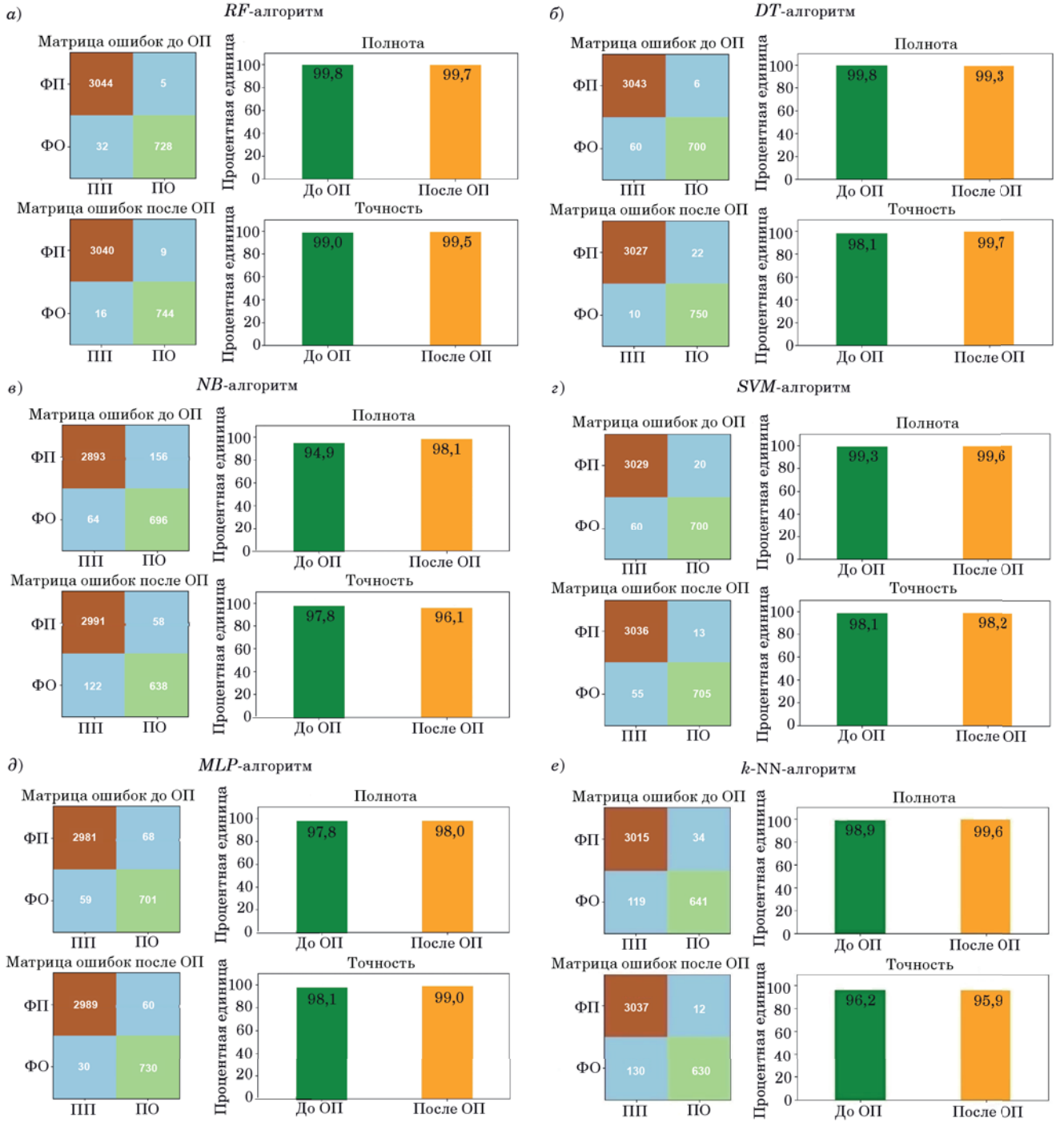
Метрика точности для всех алгоритмов представлена на рис. 2. Результаты экспериментов показывают, что точность всех алгоритмов относительно высока. Повышение точности алгоритмов также указывает на необходимость и влияние отбора признаков, который повышает точность до 1,05 %. Хотя это значение невелико, оно также частично свидетельствует об улучшении модели. Эта разница была обусловлена многими факторами. Одним из важных факторов является обучающий набор данных. Если этот набор данных достаточно велик и точен, алгоритмы будут более эффективными. Тогда эта разница станет более очевидной. Здесь мы игнорируем статистическую погрешность, поскольку оба случая выполняются в абсолютно одинаковых условиях. Производительность алгоритмов *k-NN, DT* и *MLP* является многообещающей. Их максимальная точность доходит до 99,34 %. Таким образом, можно предварительно сделать вывод, что рабочие модели стабильны, а качество набора данных хорошее.



■ ***Рис. 2.*** Сравнение алгоритмов по показателю точность (*accuracy*)
■ ***Fig. 2.*** Comparison of algorithms in terms of accuracy

Анализ на основе полноты и точности (precision)

На рисунке 3 показаны флуктуации индексов *TP*, *TN*, *FP*, *FN* в матрице ошибок каждого алгоритма; изменения соответствующих метрик, таких как полнота, точность при использовании всех признаков PE-заголовка и лучших признаков PE-заголовка. Результат наглядно показы-



■ ***Рис. 3.*** Матрица ошибок, полнота (*recall*) и точность (*precision*) алгоритмов случайный лес (*а*), дерево принятия решений (*б*), наивный байесовский алгоритм (*в*), метод опорных векторов (*г*), многослойный перцептрон (*д*), метод *k*-ближайших соседей (*е*) до и после использования отбора признаков: *ОП* — отбор признаков; *ПО* — прогнозированный отрицательный; *ПП* — прогнозированный положительный; *ФО* — фактический отрицательный; *ФП* — фактический положительный

■ ***Fig. 3.*** Confusion matrix, recall and precision of the algorithms Random Forest (*а*), Decision Tree (*б*), Naive Bayes (*в*), Support Vector Machine (*г*), Multilayer Perceptron (*д*), k-Nearest Neighbors (*е*) before and after using the feature selection

вает увеличение и уменьшение этих индексов в матрице ошибок. Большинство алгоритмов увеличивают показатель *TP*, чтобы продемонстрировать эффективность лучшего набора признаков PE-заголовка для обнаружения большего количества вредоносных программ. Для *FN* большинство значений немного уменьшаются или уменьшаются вдвое (для алгоритмов *NB, SVM* и *k-NN*) (рис. 3, *в*, *г*, *е*). Этот результат показывает, что выбор признаков помогает этим алгоритмам существенно снизить количество ошибочных прогнозов.

Индекс *FP* также значительно снизился в алгоритмах *RF*, *DT* и *MLP* (рис. 3, *а*, *б*, *д*), но почти вдвое (с 64 до 122 образцов, рис. 3, *в*) увеличился при использовании лучшего набора признаков для алгоритма *NB*. Это означает, что вычисление условной вероятности для некоторых признаков может привести к неправильному прогнозированию незараженных файлов во вредоносные. Однако в сочетании с индексом *TN* при использовании лучших признаков PE-заголовка эти алгоритмы значительно улучшаются, позволяя избежать путаницы между вредоносными и обычными файлами.

Кроме того, *RF* и *DT* можно рассматривать как два алгоритма с очень высокой вероятностью прогнозирования файла как вредоносного ПО с точностью 99,5 и 99,7 % соответственно (рис. 3, *а*, *б*). Мало того, *RF* и *k-NN* (рис. 3, *а*, *е*) имеют впечатляющий коэффициент полноты — 99,7 и 99,6 % соответственно. Это означает, что вероятность обнаружения реального вредоносного ПО с помощью *RF* и *k-NN* очень высока или вероятность пропуска реального вредоносного ПО с *RF*, *k-NN* очень мала.

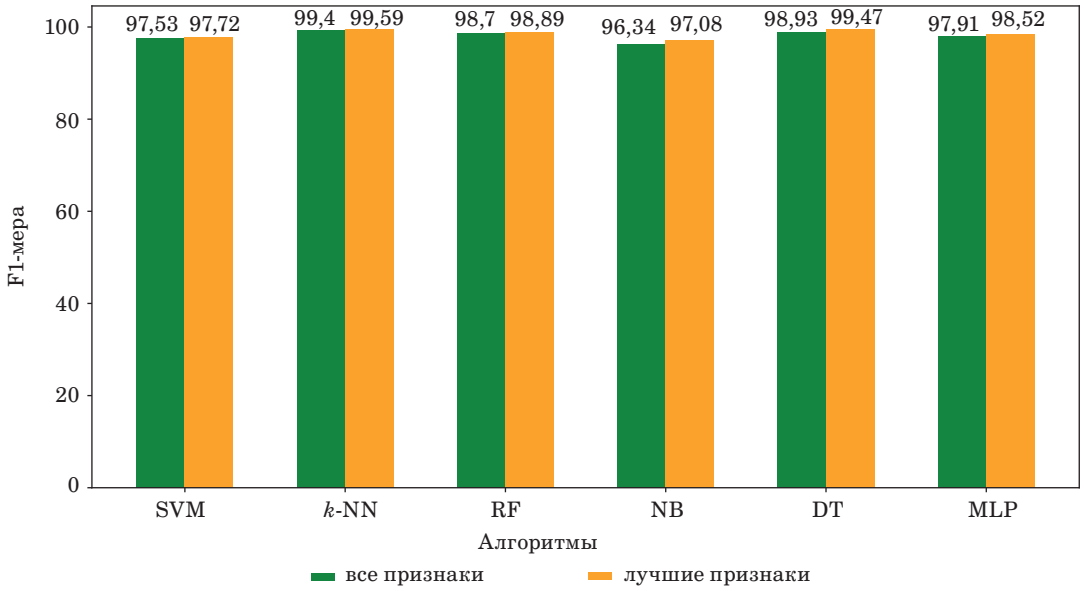
Анализ на основе F1-меры

Как известно, использование только точности или полноты метрик не позволяет правильно оценить качество модели. Поэтому для оценки вышеуказанных моделей следует использовать метрику F1-меру. На рис. 4 показано значение F1-меры всех алгоритмов в случае использования лучших признаков PE-заголовка и в случае без выбора этих признаков. Мы видим, что это значение улучшается благодаря лучшим признакам. Это означает, что лучший набор признаков PE-заголовка играет важную роль в обнаружении вредоносного ПО, поскольку чем выше F1-мера, тем лучше классификатор.

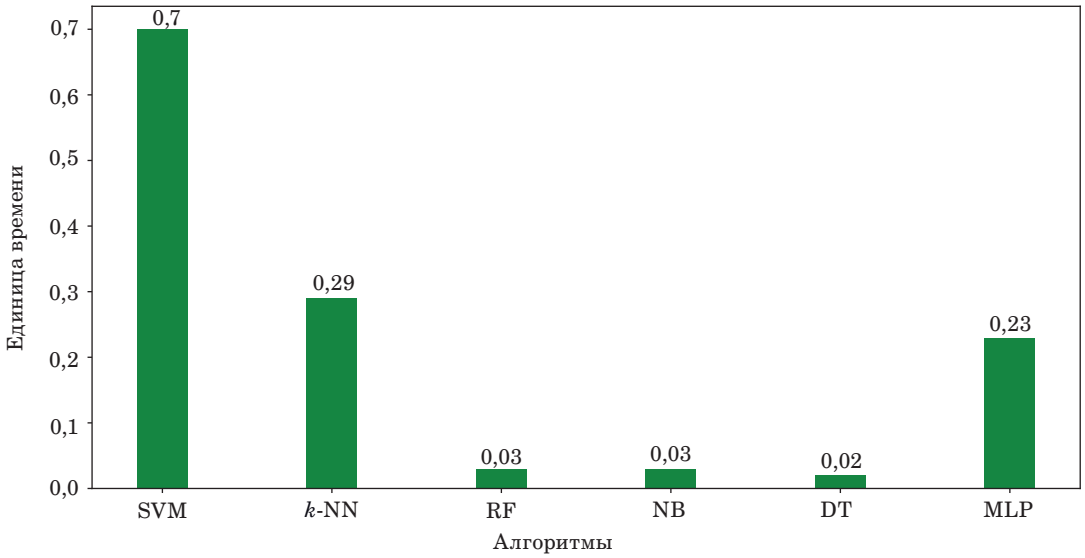
Из анализа на основе точности и полноты легко увидеть, что F1-мера для *DT* и *RF* будет очень высока. Кроме того, алгоритм *k-NN* также имеет высокую F1-меру, которая доказывает, что *k-NN* может быть полностью использован для обнаружения вредоносных программ.

В дополнение к вышеперечисленным метрикам время обработки также является важной метрикой, используемой для сравнения алгоритмов. На рис. 5 показано время обработки всех алгоритмов.

Мы видим, что *RF*, *NB* и *DT* являются алгоритмами с быстрым временем обработки. При высокой скорости обработки данных современных компьютеров обработать структуры деревьев или вычислить значения вероятности для составления прогнозов можно со скоростью



■ ***Рис. 4.*** F1-мера алгоритмов до и после отбора лучших признаков
■ ***Fig. 4.*** F1-score of the algorithms before and after using the best feature selection

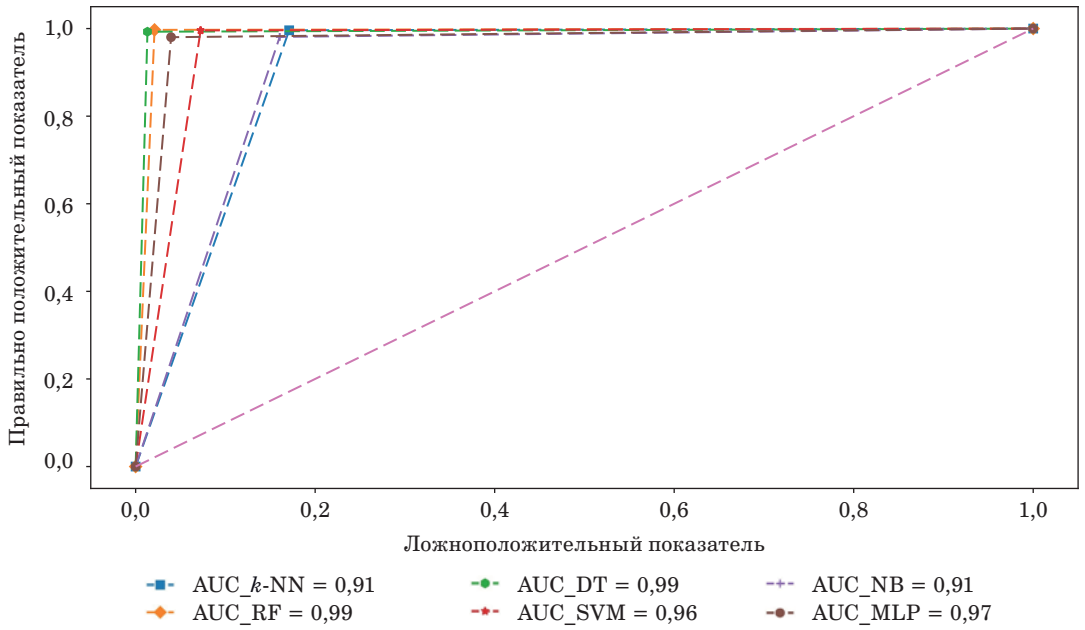


■ ***Рис. 5.*** Относительное время обработки алгоритмов
■ ***Fig. 5.*** Relative processing time of the algorithms

0,02–0,03 с. Алгоритмы *SVM*, *k-NN* и *MLP* показывают более длительное время обработки данных, поскольку оно в значительной степени зависит от природы алгоритма. В частности, вычисление расстояния между несколькими соседями (в *k-NN*), поиск суперплоскости (в *SVM*) или построение нейронной сети с большим количеством скрытых слоев (в *MLP*) для составления прогнозов занимают много времени. Кроме того, скорость процессора также частично влияет на время обработки этих алгоритмов. Поэтому экспериментальные результаты времени обработки являются относительными.

ROC-кривая всех алгоритмов показана на рис. 6. Следует отметить, что:

— чем ближе кривая проходит по левой границе, а затем идет вдоль верхней границы пространства ROC, тем точнее будет результат прогнозирования. Поэтому, чем ближе кривая приближается к диагонали 45 градусов в про-



■ ***Рис. 6.*** ROC-кривая
■ ***Fig. 6.*** ROC-curve

странстве ROC, тем менее точным является прогнозирование;

— площадь под кривой *AUC* (*Area Under the Curve*), ограниченная пространством ROC, является мерой точности прогнозирования, например, 1 — оптимально, 0,5 — плохо. Эта область является показателем хорошей или плохой дискриминации.

Основываясь на свойствах ROC-кривой в сочетании с результатами эксперимента, можно показать наилучший алгоритм прогнозирования на основе значения левой верхней границы и площади под ROC-кривой — *AUC*. Мы видим, что *RF* и *DT* являются лучшими алгоритмами для эффективного прогнозирования, поскольку их ROC-кривые проходят близко к левой верхней границе, и значение *AUC* этих алгоритмов также очень высокое (0,99/1).

Заключение

Кибератаки с использованием вредоносных программ становятся все более популярными, что делает обнаружение вредоносных программ необходимым требованием, особенно в контексте того, что вредоносные программы создаются все умнее, разнообразнее.

Среди популярных ОС именно Windows подвергается наиболее мощным и распространенным атакам вредоносными программами. Стоит отметить, что все приложения, работающие под ОС Windows, имеют формат файла PE-заголовка. В данной работе мы показываем результаты экспериментальных исследований по обнаружению вредоносных программ в ОС Windows с помощью алгоритмов машинного обучения, основанных на признаках PE-заголовка. Исходя из эксперимента, изложенного выше, и полученных результатов, делаем вывод, что применение алгоритмов машинного обучения для обнаружения вредоносных программ вполне возможно и дает хорошие результаты.

Мы протестировали шесть различных алгоритмов: случайный лес (*RF*), дерево принятия решений (*DT*), наивный байесовский алгоритм (*NB*), метод опорных векторов (*SVM*), многослойный перцептрон (*MLP*), метод *k*-ближайших соседей (*k-NN*) — с большим набором данных. Результаты показали, что:

— алгоритмы *RF*, *DT*, *k-NN* и *MLP* могут обнаруживать вредоносные образцы с очень высокой точностью (>98 %). Однако *MLP* и *k-NN* могут не подойти, поскольку время обработки значительно больше, чем у алгоритмов *DT* и *RF*;

— у *RF* очень высокие показатели точности и полноты. Этот факт доказывает, что *RF*-алгоритм очень хорошо подходит для обнаружения вредоносного ПО на ОС Windows.

В дополнение можно также рассмотреть возможность использования алгоритма *NB* в качестве альтернативы. Хотя точность *NB* не так высока, как *DT*, *k-NN* или *MLP*, она достаточно высокая (> 96 %). Кроме того, *NB* считается простым алгоритмом для реализации и имеет быстрое время обработки.

Мы также отметили лучшие признаки, позволяющие идентифицировать вредоносные файлы с высокой точностью и эффективностью. Однако в этой статье мы использовали только вредоносное ПО для архитектуры x86-64, поэтому результаты, полученные в эмпирическом процессе, могут отличаться для набора данных, содержащего вредоносное ПО для архитектуры x64.

Следует подчеркнуть, что многие категории вирусного ПО, нацеленного на ОС Windows, не имеют PE-заголовка. Следовательно, необходимо найти более комплексное новое решение для обнаружения многих типов вредоносных ПО в ОС Windows.

В будущих работах мы сосредоточимся на применении алгоритмов глубокого обучения для обнаружения вредоносных программ и их классификации.

Литература

1. **Pandey A. K., Tripathi A. K., Kapil G., Singh V., Khan M. W., Agrawal A., Kumar R., Khan R. A.** *Trends in Malware Attacks: Identification and Mitigation Strategies*. Critical Concepts, Standards, and Techniques in Cyber Forensics. IGI Global, 2020. Pp. 47–60. doi:10.4018/978-1-7998-1558-7.ch004
2. **Alrzini J., Pennington D.** A review of polymorphic malware detection techniques. *Intern. Conf. on Interdisciplinary Computer Science and Engineering (ICICSE2020)*, 2020. https://pureportal.strath.ac.uk/en/publications/a-review-of-polymorphic-malware-detection-techniques (дата обращения: 29.06.2022).
3. **Mohanta A., Saldanha A.** *Persistence Mechanisms*. Malware Analysis and Detection Engineering. Apress, Berkeley, CA, 2020. Pp. 213–236. doi:10.1007/978-1-4842-6193-4_8
4. **Afianian A., Niksefat S., Sadeghiyan B., Baptiste D.** Malware dynamic analysis evasion techniques: A survey. *ACM Computing Surveys (CSUR)*, 2019, vol. 52, no. 6, pp. 1–28. doi:10.1145/3365001
5. **Alaeiyan M., Parsa S., Conti M.** Analysis and classification of context-based malware behavior. *Computer Communications*, 2019, vol. 136, pp. 76–90. doi:10.1016/j.comcom.2019.01.003
6. **Sikorski M., Honig A.** *Practical Malware Analysis: The Hands-on Guide to Dissecting Malicious Soft-*

ware. No Starch Press, 2012. 766 p. doi:10.1016/j.cose.2012.05.004
7. **Ucci D., Aniello L., Baldoni R.** Survey of machine learning techniques for malware analysis. *Computers & Security*, 2019, vol. 81, pp. 123–147. doi:10.1016/j.cose.2018.11.001
8. **Shalaginov A., Banin S., Dehghantanha A., Franke K.** *Machine learning aided static malware analysis: A survey and tutorial*. Cyber Threat Intelligence. Springer, Cham, 2018. Pp. 7–45. doi:10.1007/978-3-319-73951-9_2
9. **Chowdhury M., Rahman A., Islam R.** Malware analysis and detection using data mining and machine learning classification. *Intern. Conf. on Applications and Techniques in Cyber Security and Intelligence*. Edizioni della Normale, Cham, 2017, pp. 266–274. doi:10.1007/978-3-319-67071-3_33
10. **Singh A., Thakur N., Sharma A.** A review of supervised machine learning algorithms. *2016 3rd Intern. Conf. on Computing for Sustainable Global Development (INDIACom)*, IEEE, 2016, pp. 1310–1315.
11. **Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J.** Scikit learn: Machine learning in Python. *The Journal of Machine Learning Research*, 2011, vol. 12, pp. 2825–2830.
12. **Gavrilut D., Cimpoesu M., Anton D., Ciortuz L.** Malware detection using machine learning. *2009 Intern. Multiconf. on Computer Science and Information Technology*, IEEE, 2009, pp. 735–741. doi:10.1109/imcsit.2009.5352759
13. **Rathore H., Agarwal S., Sahay S. K., Sewak M.** Malware detection using machine learning and deep learning. *Intern. Conf. on Big Data Analytics*, Springer, Cham, 2018, pp. 402–411. doi:10.1007/978-3-030-04780-1_28
14. **Bae S. I., Lee G. B., Im E. G.** Ransomware detection using machine learning algorithms. *Concurrency and Computation: Practice and Experience*, IEEE, 2020, vol. 32, no. 18, e.5422. doi.org/10.1002/cpe.5422
15. **Jerlin M. A., Marimuthu K.** A new malware detection system using machine learning techniques for API call sequences. *Journal of Applied Security Research*, 2018, vol. 13, no. 1, pp. 45–62. doi:10.1080/19361610.2018.1387734
16. **Ravi C., Manoharan R.** Malware detection using Windows API sequence and machine learning. *International Journal of Computer Applications*, 2012, vol. 43, no. 17, pp. 12–16. doi:10.5120/6194-8715
17. **Chaudhary S., Garg A.** A machine learning technique to detect behavior based malware. *2020 10th Intern. Conf. on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, 2020, pp. 655–659. doi:10.1109/confluence47617.2020.9058173
18. **Xu Z., Ray S., Subramanyan P., Malik S.** Malware detection using machine learning based analysis of virtual memory access patterns. *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, IEEE, 2017, pp. 169–174. doi:10.23919/date.2017.7926977
19. **Santos I., Penya Y. K., Devesa J., Bringas P. G.** N-grams-based file signatures for malware detection. *Proc. of the 11th Intern. Conf. on Enterprise Information Systems (ICEIS (2))*, 2009, vol. 9, pp. 317–320. doi:10.5220/0001863603170320
20. **Bai J., Wang J., Zou G.** A malware detection scheme based on mining format information. *The Scientific World Journal*, vol. 2014, pp. 1–12. doi:10.1155/2014/260905
21. **Kim S.** *PE header Analysis for Malware Detection*. Master Thesis, San Jose State University, 2018. 50 p.
22. **Kumar A., Kuppusamy K. S., Aghila G.** A learning model to detect maliciousness of portable executable using integrated feature set. *Journal of King Saud University — Computer and Information Sciences*, 2019, vol. 31, no. 2, pp. 252–265. doi:10.1016/j.jksuci.2017.01.003
23. **Webster G. D., Kolosnjaji B., Pentz C. V., Kirsch J., Hanif Z. D., Zarras A., Eckert C.** Finding the needle: A study of the PE32 rich header and respective malware triage. *Intern. Conf. on Detection of Intrusions and Malware, and Vulnerability Assessment*, Springer, Cham, 2017, pp. 119–138. doi:10.1007/978-3-319-60876-1_6

UDC 004.056.5
doi:10.31799/1684-8853-2022-4-44-57

Applying machine learning algorithms for PE-header-based malware detection on the Windows operating system

D. T. Le[a], PhD., Lecturer, orcid.org/0000-0003-3735-0314, letranduc@dut.udn.vn
M. H. Pham[b], Engineer, orcid.org/0000-0002-2250-9428
T. D. Dinh[c], PhD., Lecturer, orcid.org/0000-0002-9993-9792
H. P. Do[d], M. Sc., Lecturer, orcid.org/0000-0003-0645-0021
[a]The University of Danang — University of Science and Technology, Information Technology Faculty, 54 Nguyen Luong Bang, 550000, Da Nang, Vietnam
[b]R&D Department, Viettel Business Solutions Corporation, Vietnam
[c]Posts and Telecommunications Institute of Technology, 122 Hoang Quoc Viet, Hanoi, Vietnam
[d]Danang Architecture University, 566 Nui Thanh, Danang, Vietnam

Introduction: The rapid growth of malware and its malicious use result in significant financial losses for various organizations. Many researchers are interested in applying machine learning methods to solve the problem of malware detection. Nevertheless, because

of the diversity of algorithms, each machine learning algorithm has its advantages and disadvantages for a given situation. **Purpose:** To apply machine learning for malware detection in the Windows operating system using Portable Executable header; to compare six different machine learning algorithms based on several criteria. **Results:** The comparison of various algorithms, including such classifiers as Random Forest, Decision Tree, Naive Bayes, Support Vector Machine, Multilayer Perceptron, k-Nearest Neighbors algorithm with a large dataset shows that some algorithms such as Random Forest, Decision Tree, k-Nearest Neighbors, and Multilayer Perceptron can detect malware with very high accuracy (> 98%). The Random Forest algorithm is especially well suited for Windows OS malware detection. At the same time, Naive Bayes classifier also has a high accuracy rate (> 96%) and fast processing time. Therefore, we may consider using Naive Bayes as an alternative.

Keywords — malware, machine learning algorithms, PE-header, Windows.

For citation: Le D. T., Pham M. H., Dinh T. D., Do H. P. Applying machine learning algorithms for PE-header-based malware detection on the Windows operating system. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 44–57 (In Russian). doi:10.31799/1684-8853-2022-4-44-57

Reference

1. Pandey A. K., Tripathi A. K., Kapil G., Singh V., Khan M. W., Agrawal A., Kumar R., Khan R. A. *Trends in Malware Attacks: Identification and Mitigation Strategies*. In: *Critical Concepts, Standards, and Techniques in Cyber Forensics*. IGI Global, 2020. Pp. 47–60. doi:10.4018/978-1-7998-1558-7.ch004
2. Alrzini J., Pennington D. A review of polymorphic malware detection techniques. *Intern. Conf. on Interdisciplinary Computer Science and Engineering (ICICSE2020)*, 2020. Available at: https://pureportal.strath.ac.uk/en/publications/a-review-of-polymorphic-malware-detection-techniques (accessed 29 June 2022).
3. Mohanta A., Saldanha A. *Persistence Mechanisms*. In: *Malware Analysis and Detection Engineering*. Apress, Berkeley, CA, 2020. Pp. 213–236. doi:10.1007/978-1-4842-6193-4_8
4. Afianian A., Niksefat S., Sadeghiyan B., Baptiste D. Malware dynamic analysis evasion techniques: A survey. *ACM Computing Surveys (CSUR)*, 2019, vol. 52, no. 6, pp. 1–28. doi:10.1145/3365001
5. Alaeiyan M., Parsa S., Conti M. Analysis and classification of context-based malware behavior. *Computer Communications*, 2019, vol. 136. pp. 76–90. doi:10.1016/j.comcom.2019.01.003
6. Sikorski M., Honig A. *Practical Malware Analysis: The Hands-on Guide to Dissecting Malicious Software*. No Starch Press, 2012. 766 p. doi:10.1016/j.cose.2012.05.004
7. Ucci D., Aniello L., Baldoni R. Survey of machine learning techniques for malware analysis. *Computers & Security*, 2019, vol. 81, pp. 123–147. doi:10.1016/j.cose.2018.11.001
8. Shalaginov A., Banin S., Dehghantanha A., Franke K. *Machine learning aided static malware analysis: A survey and tutorial*. In: *Cyber Threat Intelligence*. Springer, Cham, 2018. Pp. 7–45. doi:10.1007/978-3-319-73951-9_2
9. Chowdhury M., Rahman A., Islam R. Malware analysis and detection using data mining and machine learning classification. *Intern. Conf. on Applications and Techniques in Cyber Security and Intelligence*. Edizioni della Normale, Cham, 2017, pp. 266–274. doi:10.1007/978-3-319-67071-3_33
10. Singh A., Thakur N., Sharma A. A review of supervised machine learning algorithms. *2016 3rd Intern. Conf. on Computing for Sustainable Global Development (INDIACom)*, IEEE, 2016, pp. 1310–1315.
11. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J. Scikit learn: Machine learning in Python. *The Journal of Machine Learning Research*, 2011, vol. 12, pp. 2825–2830.
12. Gavrilut D., Cimpoesu M., Anton D., Ciortuz L. Malware detection using machine learning. *2009 Intern. Multiconf. on Computer Science and Information Technology*, IEEE, 2009, pp. 735–741. doi:10.1109/imcsit.2009.5352759
13. Rathore H., Agarwal S., Sahay S. K., Sewak M. Malware detection using machine learning and deep learning. *Intern. Conf. on Big Data Analytics*, Springer, Cham, 2018, pp. 402–411. doi:10.1007/978-3-030-04780-1_28
14. Bae S. I., Lee G. B., Im E. G. Ransomware detection using machine learning algorithms. *Concurrency and Computation: Practice and Experience*, IEEE, 2020, vol. 32, no. 18, e.5422. doi.org/10.1002/cpe.5422
15. Jerlin M. A., Marimuthu K. A new malware detection system using machine learning techniques for API call sequences. *Journal of Applied Security Research*, 2018, vol. 13, no. 1, pp. 45–62. doi:10.1080/19361610.2018.1387734
16. Ravi C., Manoharan R. Malware detection using Windows API sequence and machine learning. *International Journal of Computer Applications*, 2012, vol. 43, no. 17, pp. 12–16. doi:10.5120/6194-8715
17. Chaudhary S., Garg A. A machine learning technique to detect behavior based malware. *2020 10th Intern. Conf. on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, 2020, pp. 655–659. doi:10.1109/confluence47617.2020.9058173
18. Xu Z., Ray S., Subramanyan P., Malik S. Malware detection using machine learning based analysis of virtual memory access patterns. *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, IEEE, 2017, pp. 169–174. doi:10.23919/date.2017.7926977
19. Santos I., Penya Y. K., Devesa J., Bringas P. G. N-grams-based file signatures for malware detection. *Proc. of the 11th Intern. Conf. on Enterprise Information Systems (ICEIS (2))*, 2009, vol. 9, pp. 317–320. doi:10.5220/0001863603170320
20. Bai J., Wang J., Zou G. A malware detection scheme based on mining format information. *The Scientific World Journal*, vol. 2014, pp. 1–12. doi:10.1155/2014/260905
21. Kim S. *PE Header Analysis for Malware Detection*. Master Thesis, San Jose State University, 2018. 50 p.
22. Kumar A., Kuppusamy K. S., Aghila G. A learning model to detect maliciousness of portable executable using integrated feature set. *Journal of King Saud University — Computer and Information Sciences*, 2019, vol. 31, no. 2, pp. 252–265. doi:10.1016/j.jksuci.2017.01.003
23. Webster G. D., Kolosnjaji B., Pentz C. V., Kirsch J., Hanif Z. D., Zarras A., Eckert C. Finding the needle: A study of the PE32 rich header and respective malware triage. *Intern. Conf. on Detection of Intrusions and Malware, and Vulnerability Assessment*, Springer, Cham, 2017, pp. 119–138. doi:10.1007/978-3-319-60876-1_6

УДК 004.946
doi:10.31799/1684-8853-2022-4-58-67

Поиск закономерностей на ЭЭС при симуляции тревожно-фобической ситуации в среде виртуальной реальности

А. Ю. Тычков[а], *доктор техн. наук, профессор, orcid.org/0000-0002-2354-2895, tychkov-a@mail.ru*
Д. С. Чернышов[а], *инженер-программист, orcid.org/0000-0002-1604-7001*
П. П. Чураков[а], *доктор техн. наук, профессор, orcid.org/0000-0002-5918-732X*
З. М. Юлдашев[б], *доктор техн. наук, профессор, orcid.org/0000-0003-1075-3420*
Н. С. Бофанова[а], *канд. мед. наук, доцент, orcid.org/0000-0002-5455-4987*
А. К. Алимурадов[а], *канд. техн. наук, доцент, orcid.org/0000-0002-5133-2713*
В. Н. Горбунов[а], *канд. экон. наук, доцент, orcid.org/0000-0003-0195-0772*
Р. В. Золотарев[а], *инженер-программист, orcid.org/0000-0002-5614-6557*
М. С. Никитин[а], *инженер-программист, orcid.org/0000-0002-5708-3391*
[а]*Пензенский государственный университет, Красная ул., 40, Пенза, 440000, РФ*
[б]*Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», Профессора Попова ул., 5, Санкт-Петербург, 197022, РФ*

Введение: *эффективная и оперативная постановка диагностических заключений о наличии тревожно-фобических расстройств требует совершенствования существующих и разработки новых способов диагностики и лечения пациентов, в том числе с применением технологии виртуальной реальности.* ***Цель:*** *исследовать реакцию испытуемого на стимул, вызывающий страх, противопоставленный через сцены виртуальной реальности (нахождение на высоте). Определить маркеры на электроэнцефалографическом сигнале, ассоциированные с уровнем тревоги и восприимчивости испытуемого в среде виртуальной реальности.* ***Методы:*** *испытуемые погружались в виртуальную реальность с прохождением анкетирования самооценки, регистрацией электроэнцефалографического сигнала и спектральным анализом сигналов в различных частотных диапазонах в среде статистического программирования R версии 4.1.2.* ***Результаты:*** *для проведения исследования сформирована группа испытуемых из девяти условно здоровых мужчин в возрасте от 23 до 26 лет, отрицавших хронические соматические заболевания в анамнезе, органические заболевания головного мозга. Проведено сравнение состояния покоя (эталонного значения) и состояния в среде виртуальной реальности испытуемых в момент высокой эмоциональной нагрузки (на высоте небоскреба). Полученные результаты позволяют сделать вывод, что симуляция ситуации нахождения на высоте у различных испытуемых, независимо от интенсивности проявления страха, вызывает снижение индексов альфа-, тета-, бета-ритмов и увеличение индекса дельта-ритма на электроэнцефалографическом сигнале относительно состояния покоя. Выявленные закономерности позволяют оценить уровень тревоги человека в среде виртуальной реальности.* ***Практическая значимость:*** *результаты исследований могут быть использованы при разработке новых клинических методик диагностики тревожно-фобических расстройств с применением технологии виртуальной реальности. Они позволят повысить уровень объективной оценки психоэмоционального состояния человека путем выявления объективных электрофизиологических маркеров, зарегистрированных у человека в период нахождения его в среде виртуальной реальности.*

Ключевые слова *— виртуальная реальность, тревожно-фобические расстройства, электроэнцефалографический сигнал.*

Для цитирования: Тычков А. Ю., Чернышов Д. С., Чураков П. П., Юлдашев З. М., Бофанова Н. С., Алимурадов А. К., Горбунов В. Н., Золотарев Р. В., Никитин М. С. Поиск закономерностей на ЭЭС при симуляции тревожно-фобической ситуации в среде виртуальной реальности. *Информационно-управляющие системы*, 2022, № 4, с. 58–67. doi:10.31799/1684-8853-2022-4-58-67
For citation: Tychkov A. Yu., Chernyhov D. S., Churakov P. P., Yuldashev Z. M., Bofanova N. S., Alimuradov A. K., Gorbunov V. N., Zolotarev R. V., Nikitin M. S. Search for EEG signal patterns in simulating phobic anxiety disorder situations in a VR environment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 58–67 (In Russian). doi:10.31799/1684-8853-2022-4-58-67

Введение

На современном этапе развития информационных технологий наблюдается тенденция глобальной цифровизации всех сфер жизнедеятельности человека, в том числе здравоохранения. Технология виртуальной реальности (ВР; Virtual Reality, VR) как инструмент развития «цифровой медицины» способен замещать или искажать информацию от органов чувств, изменяя и анализируя при этом ответные реакции организма, осуществляя взаимодействие организма с создаваемой виртуальной средой [1]. Технология ВР находит широкое применение в игровой индустрии, симуляторах различного назначения, системах медицинской диагностики и лечения пациентов. В настоящее время ВР является перспективным и динамично развивающимся направлением, особенно в области неврологии и психиатрии. ВР появилась как альтернатива создания виртуальных стимулов, близких к реальной жизни, что позволило человеку взаимодействовать с окружающей средой в различных ситуациях [2].

В настоящее время в связи с пандемией, а также сложной геополитической обстановкой в мире отмечается рост обращений за медицинской помощью к врачам-неврологам, врачам-психиатрам и психотерапевтам. У каждого четвертого жителя планеты возможно диагностировать тревожно-фобическое расстройство в связи с неблагоприятными условиями среды и ростом стрессовых факторов.

Для постановки диагноза и оценки тяжести состояния пациентов врачами применяется субъективный сбор анамнеза, а также интерпретация психометрических шкал, которые имеют субъективный характер. В настоящее время отсутствие объективной инструментальной диагностики тревожно-фобических расстройств может приводить к необъективной интерпретации результатов и, как следствие, неправильной постановке диагноза [3, 4].

Электроэнцефалография — важный метод изучения нейрофизиологических механизмов тревожно-фобических расстройств, который позволяет оценивать текущее функциональное состояние различных отделов коры головного мозга.

Для эффективной и оперативной постановки диагностических заключений о наличии тревожно-фобических расстройств необходимо совершенствовть существующие и разрабатывать новые способы диагностики и лечения, в том числе с применением ВР.

Виртуальная реальность в диагностике и лечении тревожно-фобических расстройств

Виртуальная реальность — это тип человеко-компьютерного интерфейса, который позволяет пользователям взаимодействовать с компьютерами в режиме реального времени и погружаться в среду, созданную компьютерами. Пользователями ВР являются более 40 млн человек в мире (2019 — 35 млн чел., 2020 — 42 млн чел.). Технология ВР прошла длинный путь от первых экспериментов в 50-х годах XX века до современных беспроводных шлемов в 20-х годах XXI века [5].

Применение ВР позволяет передавать аудиовизуальные переживания более реалистично, чем традиционные терапевтические технологии, и обеспечивает безопасную и контролируемую среду. Сценарий настраивается в соответствии с состоянием пациента таким образом, чтобы страх, радость и другие эмоции пациента могли быть полностью выражены [6].

Около 12 % населения земного шара (порядка миллиарда человек) имеют тревожно-фобические расстройства. Тревожно-фобические расстройства [7] — группа расстройств, в клинической картине которых преимущественно встречается тревога, страх конкретных ситуаций или объектов, которые не являются опасными.

Согласно классификации тревожно-фобических расстройств [7], выделяются:

1) агорафобия — страх перед посещением общественных мест, магазинов, транспорта, страх перед выходом на улицу;

2) социальные фобии — неконтролируемый страх ситуаций, которые связаны с действиями в присутствии других людей, а также страх негативной оценки со стороны других людей;

3) специфические фобии — навязчивый страх определенных животных (например, собак, кошек, пауков, змей), людей (клоунов, дантистов, врачей и пр.), окружающей среды (темных мест, гроз, высоких мест и т. п.) или ситуаций (таких как летать на самом самолете, ехать на поезде, находиться в очень замкнутом пространстве, страх лифтов).

Исследования, посвященные вопросу применения технологии ВР в лечении пациентов с тревожно-фобическими расстройствами, малочисленны [8–14], работы, в которых виртуальную среду использовали как диагностический инструмент, — единичны [13, 14].

Экспозиционная терапия, несмотря на доказанную эффективность как метода воздействия на пациентов с тревожно-фобическим расстройством, имеет ряд недостатков: данный вид терапии требует от пациента повествования о предмете или ситуации, которая вызывает страх. Этот факт приводит к низкой эмоциональной вовлеченности пациентов, потому что они, как правило, избегают любого напоминания о ситуации или объекте, который вызывает чувство неконтролируемого страха, и, как следствие, к неудовлетворительным результатам терапии.

Экспозиционная терапия тревожно-фобических расстройств с использованием технологии ВР состоит из поэтапного постепенного погружения в обстановку или ситуацию, которая вызывает страх. Со временем чувство страха становится меньше и обстановка, которая вызывала страх, становится комфортнее. Технология ВР позволяет проще погрузиться в ситуацию, потому что происходит воздействие на зрительные, слуховые, тактильные рецепторы, а не только за счет воображения человека как при работе с врачом-психотерапевтом [15].

В 2020 г. Г. Ланкон и др. [16] провели исследования в среде ВР в отношении испытуемого с тревожно-фобическим расстройством (страх перед акулами). Во время первоначальной оценки испытуемого основная проблема была связана с сильным беспокойством пребывания человека в водной среде. Когда пользователь находился

в воде, на него воздействовали образы в виде водорослей или камней, что приводило к возникновению панической атаки. Для решения данной проблемы авторы [16] погружали пользователя в сцены ВР с различными морскими млекопитающими и добивались привыкания к водной среде. Сравнение состояния испытуемого до и после применения ВР показало уменьшение сквалофобии (64,34 %), уменьшение расстройства настроения (33,33 %) и снижение тревожности.

На сегодня в области медицины для работы с тревожно-фобическими расстройствами существует множество программно-аппаратных решений, таких как C2Phobia, PSIOUS, Stim Response Virtual Reality, Virtually Better и др. Все известные программно-аппаратные комплексы ориентированы на коррекцию специфических и социальных фобий в отсутствие постоянного контроля поведения человека в режиме реального времени.

Постановка эксперимента исследования в среде ВР

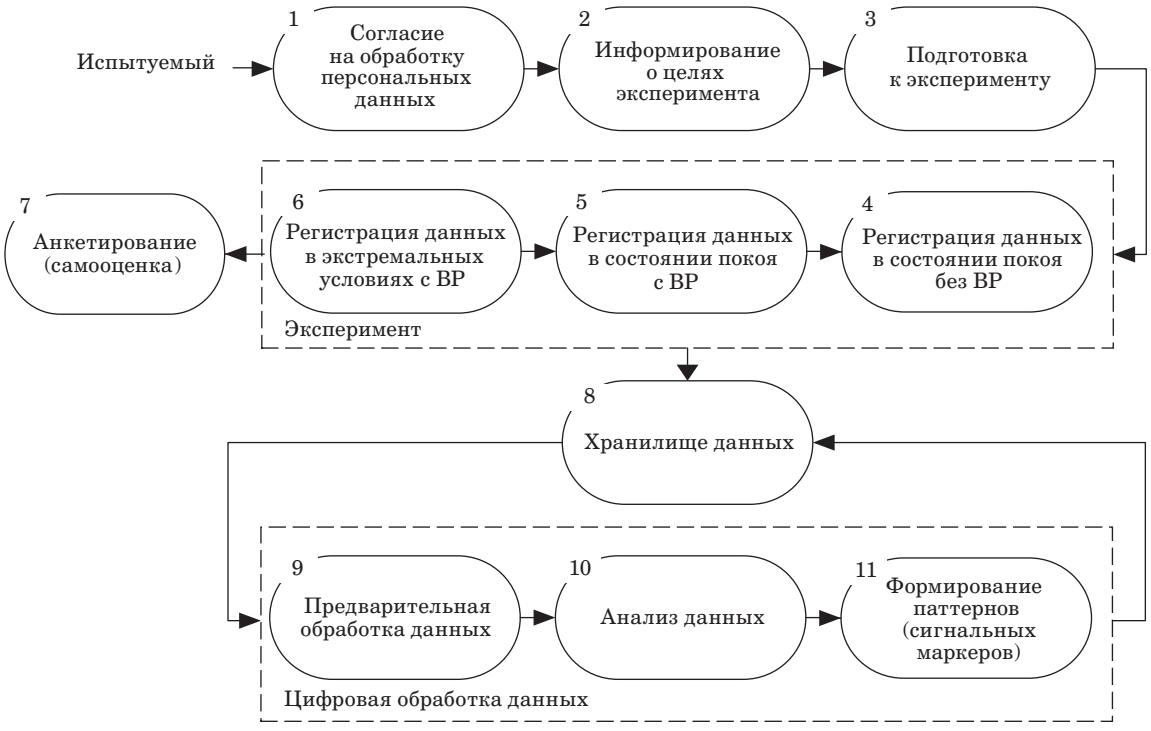
Несмотря на значительное количество готовых программно-аппаратных решений, существует необходимость в оценке состояния испытуемого на основе анализа физиологических данных.

Коллективом авторов настоящей статьи ранее разработана методика исследования тревожно-фобических расстройств в среде ВР с использованием восьмиканального электроэнцефалографа [17].

Для проведения исследования сформирована группа испытуемых из девяти условно здоровых мужчин в возрасте от 23 до 26 лет, которые отрицали хронические соматические заболевания в анамнезе, органические заболевания головного мозга. Все исследуемые дали согласие на погружение в виртуальную среду и не получали компенсации за участие. Все исследования проводились в соответствии с протоколом клинического исследования, утвержденного на заседании локального комитета по этике ФГБОУ ВО «Пензенский государственный университет» (протокол № 8 от 29 апреля 2022 г.).

Схема эксперимента показана на рис. 1.

В целях повышения эффективности реализации методики исследования тревожно-фобических расстройств первостепенное внимание уделялось вопросу оценки состояния пользователя в среде ВР, безопасности воздействия виртуальной среды на пользователя, а также оценке его виртуального опыта. По данным литературы [18–20], погружение в виртуальную среду может сопровождаться появлением ряда побочных эффектов, которые могут проявляться в виде головной боли, головокружения, чувства тошноты, наруше-



■ ***Рис. 1.*** Схема исследования тревожно-фобических расстройств в среде ВР
■ ***Fig. 1.*** Schematic for the research into phobic anxiety disorders in the VR environment

ния координации, дезориентации в пространстве (киберболезнь или симуляторное расстройство). Причина появления данных явлений — неподвижное состояние человека и передача зрительной информации о перемещении тела пользователя в виртуальном пространстве в головной мозг, возникновение ощущения иллюзии движения и перемещения в пространстве, векции.

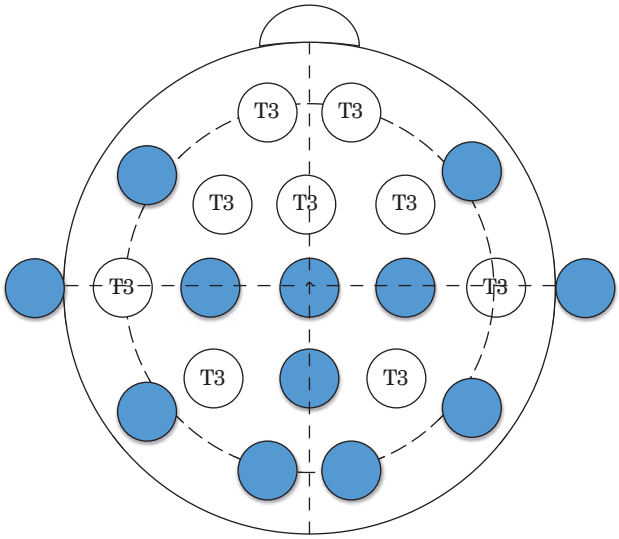
У всех исследуемых была осуществлена проверка функции зрения: все исследуемые имели нормальное зрение, у двух человек зрение до уровня нормальных показателей корректировалось с использованием контактных линз. Опыт нахождения в виртуальном пространстве имели семь участников, которые ранее не испытывали побочные эффекты нахождения в виртуальной среде. После нахождения в среде ВР всем исследуемым проведены сбор анамнеза, физикальный осмотр и неврологическое обследование врачом-неврологом, опрос по анкете SSQ с целью исключить симуляторное расстройство у пользователей, а также нейропсихологическое обследование с использованием шкалы тревоги Спилбергера — Ханина.

Для предотвращения акустической ориентации исследование проводилось в отдельном кабинете площадью 40 кв. м в условиях полной звуковой обособленности. Всем исследуемым врачом-неврологом объяснялась процедура исследования с коротким (2–3 мин) погружением в среду ВР в целях ознакомления и обучения навыкам навигации в среде виртуальной реальности (блок *3*). В этот период времени испытуемый пребывал в состоянии покоя, восстанавливал дыхание, успокаивался.

Затем (блок *4*) проводилась регистрация электроэнцефалографических сигналов (ЭЭС) в течение 5 мин в состоянии покоя с открытыми глазами. Затем ассистент надевал шлем ВР на испытуемого с продолжающейся записью ЭЭС для последующего анализа (блоки *5*, *6*). После завершения погружения в виртуальную среду все обследуемые оставались в кресле в спокойном и расслабленном состоянии еще 5 мин для контрольной записи ЭЭС, состояние также контролировал врач-невролог, проводя нейропсихологическое обследование.

Регистрация ЭЭС осуществлялась посредством электроэнцефалографа «Нейрон-спектр» по восьми отведениям. Схема расположения электродов приведена на рис. 2. В качестве последующего анализа использовались зарегистрированные ЭЭС в лобных отведениях (Fp1, Fp2). Частотное распределение ритмов ЭЭС регистрировалось в следующих диапазонах: альфа (8–12 Гц), бета (12–35 Гц), тета (4–8 Гц), дельта (0,5–4 Гц).

Электроэнцефалографический сигнал — это сигнал, регистрирующий электрическую активность головного мозга. Корковая активность основана на связи и коммуникации между нейронами, что стало возможным благодаря передаче небольших электрических сигналов, называемых электрическими импульсами.



■ ***Рис. 2.*** Схема расположения электродов
■ ***Fig. 2.*** Electrode layout

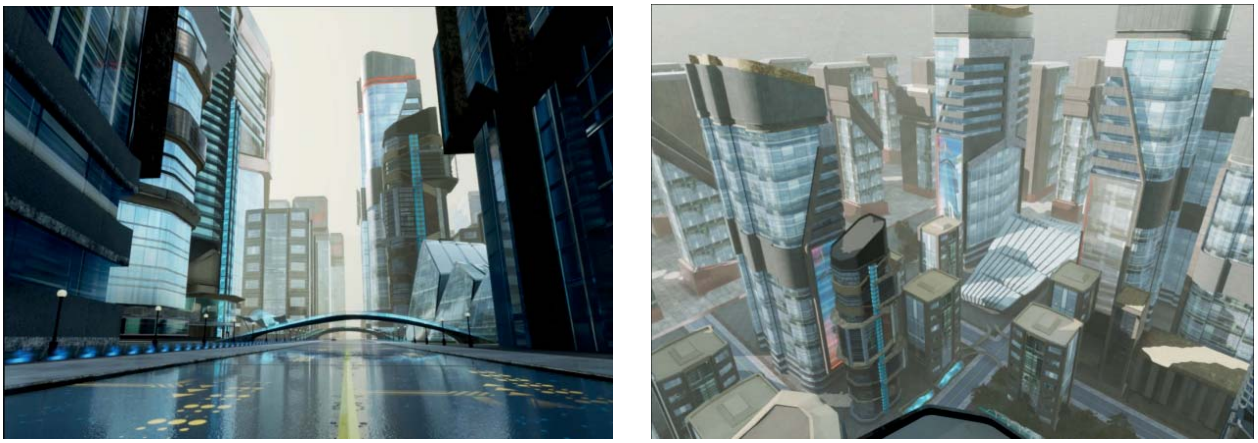
Основная задача эксперимента — продемонстрировать реакцию испытуемого на стимул, вызывающий страх, противопоставленный через сцены ВР.

После регистрации ЭЭС испытуемому предлагалось пройти анкетирование самооценки (блок *7*) для субъективного заключения об уровне страха, с которым он столкнулся в сценах ВР. Результаты регистрации заносились в хранилище данных для последующего анализа и обработки результатов измерений (блоки *8–11*).

Для демонстрации виртуального мира использовался шлем ВР — головной дисплей HTC VIVE Pro (включающий в себя два экрана AMOLED с диагональю 3.5" каждый), контроллеры. Персональный компьютер 32 ГБ оперативной памяти DDR4, графическая карта GeForce GTX 1060 6ГБ, процессор Intel Core i-3 3570. Все оборудование располагается в лаборатории «Биомедицинские и когнитивные технологии» Пензенского государственного университета.

Для проведения настоящего исследования разработана авторская сцена ВР (рис. 3) в среде моделирования Unreal Engine 4+Blender, позволяющей создать реалистичные сцены и добиться максимального эффекта погружения в виртуальную среду.

При создании сцены были использованы объекты окружения, созданные с помощью технологии фотограмметрии, что позволило добиться максимальной реалистичности сцены при малой



■ ***Рис. 3.*** Авторская сцена ВР
■ ***Fig. 3.*** Authoring VR scene

ресурсной стоимости внутренних компонентов. Для этого был выдержан единый стиль оформления и цветовое решение. Также учитывались такие факторы, как освещение: естественные и искусственные источники освещения, туман, дымка горизонта.

Пользователи погружались в среду виртуального города будущего, который представлял собой различные городские пейзажи с домами-небоскребами, улицами, водными пейзажами. Виртуальная сцена разработана от первого лица. Для обзора городского пейзажа необходимо пройти в лифт и нажать кнопку подъема. Пользователи за виртуальными пейзажами наблюдали из кабины лифта с прозрачными стенами, который поднимался с минимальной скоростью на значительную высоту. Управление лифтом осуществлялось с помощью кнопок, расположенных на джойстиках в руках пользователя, что активизировало моторную и сенсорную системы человека. При разработке сцены уделялось особое внимание тому, чтобы действия в виртуальной среде были простыми и интуитивно понятными. Зрительная информация, поступающая с помощью шлема ВР, содержала сигналы, исключающие резкие раздражающие факторы или провокационные действия. В завершение сцены двери лифта открывались, и пользователи оказывались на высоте города будущего, вместо пола лифта — узкая деревянная доска, по которой необходимо было пройти. Эта виртуальная сцена являясь неожиданной и непрогнозируемой для пользователя, вызывала чувство тревоги и беспокойства.

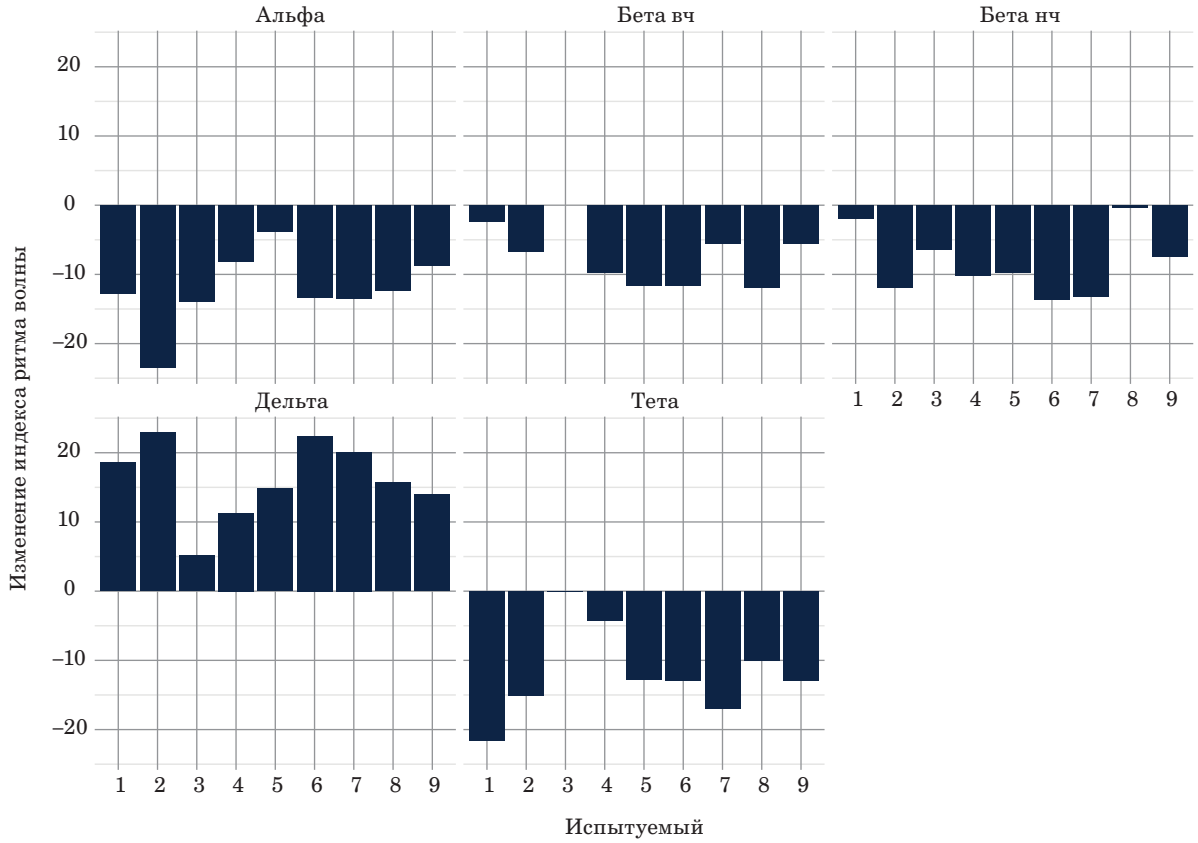
Результаты и выводы

На основании эксперимента проведен спектральный анализ изменения индекса ЭЭС в различных частотных диапазонах. Анализ произведен в среде статистического программирования R версии 4.1.2 [21] с использованием свободно распространяемой интегрированной среды для разработки (integrated development environment) на языке R — RStudio 2021.09.0+351 [22]. Для обработки и визуализации данных использованы пакеты для языка R: easyalluvial, GGally, ggplot2, parcats, readxl, scales, tidyverse [23, 24].

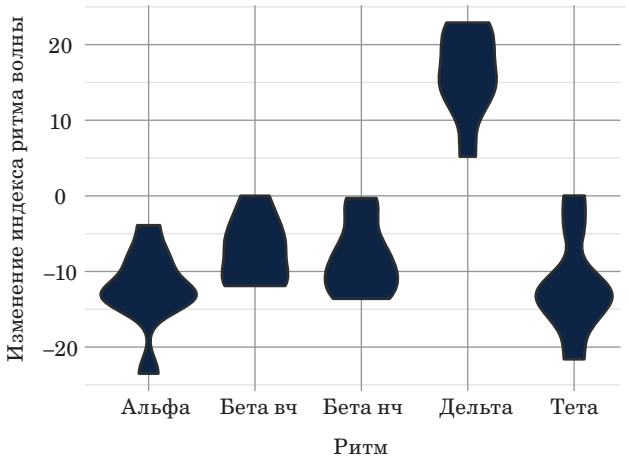
В приведенных диаграммах (рис. 4) показаны изменения индексов ритмов волн в различных частотных диапазонах. Показатели изменения выражены в абсолютных значениях. В работе сравнивается состояние покоя (эталонное значение) и состояние, в котором испытуемый находится в ВР в период высокой эмоциональной нагрузки, на высоте небоскреба. Сравнение производится относительно значений, принятых за эталонные (состояние покоя испытуемого).

Полученные результаты позволяют сделать вывод, что симуляция ситуации нахождения на высоте у различных пользователей, независимо от интенсивности проявления страха, вызывает снижение индексов альфа-, тета-, бета-ритмов и увеличение индекса дельта-ритма ЭЭС относительно состояния покоя.

Диаграмма распределения изменения (выраженного в абсолютных значениях) индексов ритмов волн (рис. 5) позволяет количественно оценить тенденции изменения индексов ритмов волн, ширина диаграммы коррелирует с частотностью. Альфа-ритм (α) снижается в пределах от 3,9 до 23,5 %. Расширение на диаграмме при $\alpha \in [-11\ \%; -16\ \%]$ показывает, что изменение индекса ритма волн в альфа-диапазоне среди испытуемых наиболее часто лежит в данных пределах. У наиболее восприимчивых к страху высоты испытуемых изменение альфа-ритма достигает максимальных значений $\Delta\alpha \in [-16\ \%;$



■ ***Рис. 4.*** Диаграммы изменения индексов ритмов волн
■ ***Fig. 4.*** Diagrams illustrating changes in wave rhythm indices

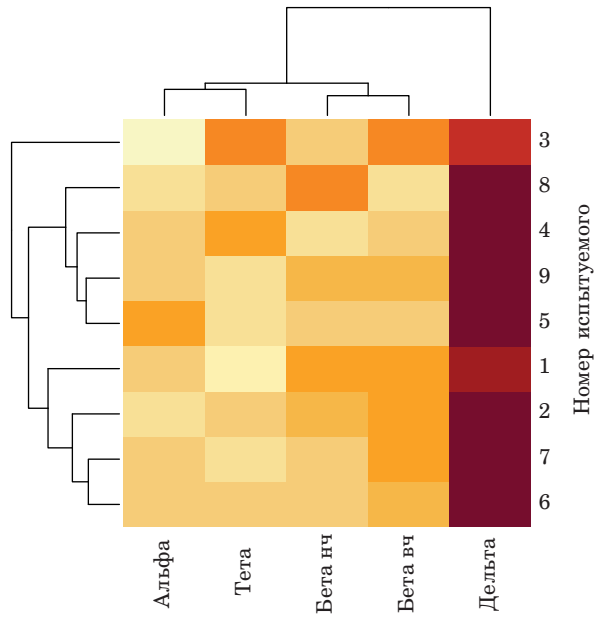


■ ***Рис. 5.*** Скрипичная диаграмма, %
■ ***Fig. 5.*** Violin plot, %

–23,5 %]. Альфа-ритм на ЭЭС указывает на процесс торможения нервной системы. Уменьшение альфа-индекса альфа-ритма в ходе исследования сигнализирует о возбуждении нервной системы испытуемых. Изменение индекса тета-ритма (θ) коррелирует с изменением индекса альфа, которое составляло [–14 %; –16 %]. Изменение индекса бета (β) равномерно распределено в пределах от десятой части процента до десяти процентов. Колебания дельта-ритма (δ) наблюдаются на ЭЭС при некоторых формах стресса. Дельта-ритм — единственный ритм, индекс которого увеличивался в ходе испытания. Испытуемые при контакте с «виртуальной высотой» испытывали чувство тревоги. У испытуемых с наименьшим проявлением чувства тревоги на высоте $\Delta\delta \in [5\ \%; 12\ \%]$, с наиболее выраженным проявлением чувства тревоги $\Delta\delta \in [12\ \%; 23\ \%]$.

Для выявления структуры в данных использовано графическое представление данных в виде тепловой карты (рис. 6), на которой каждый столбец представляет собой переменную (волны различных частотных диапазонов на ЭЭС), а каждая строка — наблюдение, интенсивностью цвета показаны величины изменений индексов ритмов волн.

Порядок строк и столбцов на тепловой карте был изменен переупорядочиванием как переменных, так и наблюдений с использованием алгоритма кластеризации: вычислено расстояние между каждой парой строк и столбцов и наблюдения упорядочены по их сходству. По значе-



■ ***Рис. 6.*** Тепловая карта изменений индексов ритмов волн у испытуемых с кластеризацией и упорядочиванием по наблюдениям и переменным

■ ***Fig. 6.*** Heat map illustrating changes in wave rhythm indices in subjects with clustering and ordering by observations and variables

ниям переменных испытуемые разделись на две группы (в 1-й группе испытуемые 1, 2, 3, 7; во 2-й группе испытуемые 3, 4, 5, 8, 9). Наблюдается сходство между изменениями ритмов волн альфа и тета, а также бета нч и бета вч.

С помощью методики Ч. Д. Спилбергера по выявлению личностной и ситуационной тревожности (адаптированной Ю. Л. Ханиным) врач определял уровень выраженности тревожности у испытуемых. Показатель ситуационной тревожности составил 46,3 ± 5,1, личностной тревожности — 54,6±5,6.

Таким образом, у всех исследуемых условно здоровых испытуемых после погружении в среду ВР уровень тревожности как ситуационной, так и личностной был выше нормальных показателей.

Полученные результаты позволяют предположить гипотезу о наличии объективных электрофизиологических маркеров, которые ассоциированы с уровнем тревоги у пользователя. Впоследствии данные показатели можно использовать в качестве диагностики тревожно-фобических расстройств у человека, находящегося в среде ВР. Это позволит отойти от субъективной оценки по шкалам психоэмоционального состояния человека (ответы на вопросы могут быть даны необъективно), заменив их оценкой объективных электрофизиологических маркеров, зарегистрированных в среде ВР.

Заключение

Проведенное исследование — одно из первых в оценке влияния технологий ВР на психоэмоциональное состояние человека. Определены объективные электрофизиологические маркеры, ассоциированные с уровнем тревоги у пользователя, которые подтверждают наличие специфических отклонений функционального состояния коры головного мозга у человека с чувством тревоги при нахождении на высоте в виртуальной среде.

В последующих исследованиях необходимо изучить вопрос влияния эффекта привыкания пользователя при нахождении в среде ВР, происходят ли адаптационные изменения организма, которые формируют виртуальный опыт, изменяющий индивидуальные характеристики взаимодействия с виртуальной средой.

Финансовая поддержка

Работа выполнена при поддержке гранта Фонда содействия развитию малых форм предприятий в научно-технической сфере, № 0015902.

Литература

1. **Карпов О. Э., Даминов В. Д., Новак Э. В., Мухаметова Д. А., Слепнева Н. И.** Технологии виртуальной реальности в медицинской реабилитации, как пример современной информатизации здравоохранения, *Вестник Национального медико-хирургического центра им. Н. И. Пирогова*, 2020, т. 15, № 1, с. 89–98.
2. **Pereira J. S., Faêda L. M., Coelho A. M.** Evolution of VRET to assist in the treatment of phobias: A systematic review. *22nd Symp. on Virtual and Augmented Reality*, 2020, pp. 386–390. doi:10.1109/SVR51698.2020.00064
3. **Li Y., Chiu P., Yeh S., Zhou C., Chen J.** Effects of virtual reality and augmented reality on induced anxiety. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, vol. 26, no. 7, pp. 1–8. doi:10.1109/ES.2017.28
4. **Krohn S., Tromp J., Quinque E. M., Belger J., Klotzsche F., Gaebler M., Thöne-Otto A., Finke C.** Multidimensional assessment of virtual reality applications in clinical neuropsychology: The “VR-check” protocol. *Intern. Conf. on Virtual Rehabilitation*, 2019, 12 p. doi:10.1109/ICVR46560.2019.8994590
5. **Tychkov A. Yu., Chernyshov D. S., Bofanova N. S., Alimuradov A. K., Ovchinnikov D. L., Sotnikov A. M.**

Virtual reality implementation for assessment and treatment of phobic anxiety disorders. *5th Scientific School Dynamics of Complex Networks and their Applications*, 2021, pp. 202–205. doi:10.1109/DCNA53427.2021.9587127

6. **Бофанова Н. С., Буланов А. А., Яворский А. С., Алехина Е. В.** Технология виртуальной реальности как современное направление в реабилитации пациентов с фантомной болью. *Российский журнал боли*, 2021, № 19(2), с. 33–37. doi:10.17116/pain20211902133
7. **Караваева Т. А., Васильева А. В., Мизинова Е. Б., Полторак С. В., Ляшковская С. В.** Клинические рекомендации «Тревожно-фобические расстройства F40». М., Российское общество психиатров, 2015. 50 с.
8. **Botella C., Fernandez-Alvarez J., Guillén V., García-Palacios A., Baños R.** Recent progress in virtual reality exposure therapy for phobias: A systematic review. *Current Psychiatry Reports*, 2017, vol. 19, iss. 7, 13 p. doi:10.1007/s11920-017-0788-4
9. **Indovina I., Conti A., Lacquaniti F., Staab J. P., Passamonti L., Toschi N.** Lower functional connectivity in vestibular-limbic networks in individuals with subclinical agoraphobia. *Frontiers in Neurology*, 2019, vol. 10, 13 p. doi:10.3389/fneur.2019.00874
10. **Hasani J., Rasti A.** The comparison of implicit and explicit memory bias to negative information processing in adolescents with high and low trait anxiety. *Journal of Psychology*, 2010, vol. 4, pp. 11–22.
11. **Gaur S., Panjwani U., Kumar B.** EEG brain wave dynamics: A systematic review and meta analysis on effect of yoga on mind relaxation. *J Biomed Res Environ Sci*, 2020, vol. 1, iss 7, pp. 353–362. doi:10.37871/jbres1166
12. **Freeman D., Reeve S., Robinson A., Ethlers A., Clark D., Spanlang B., Slater M.** Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 2017, vol. 47, iss. 14, pp. 2393–2400. doi:10.1017/S003329171700040X
13. **Jang D. P., Ku J. H., Choi Y. H., Wiederhold B. K., Nam S. W., Kim In Y., Kim Sun I.** The development of virtual reality therapy (VRT) system for the treatment of acrophobia and therapeutic case. *IEEE Trans Inf Technol Biomed*, 2002, pp. 213–220. doi:10.1109/titb.2002.802374
14. **Jadhav L., Parekh K., Gupta V., Chandak S.** Using virtual reality for therapeutic treatment of phobia. *IEEE Intern. Conf. on Convergence to Digital World — Quo Vadis*, 2020, 8 p. doi:10.1109/ICCDW45521.2020.9318727
15. **Бофанова Н. С., Петрова Е. В., Калистратов В. Б., Нестеренко Е. Н., Чиж Д. И.** Применение технологии виртуальной реальности для лечения болевого синдрома у детей. *Ульяновский медико-биологический журнал*, 2020, № 4, с. 19–29. doi:10.34014/2227-1848-2020-4-19-29
16. **Malbos E., Burgess G., Lançon C.** Virtual reality and fear of shark attack: A case study for the treatment of squalophobia. *Clinical Case Studies*, vol. 19, pp. 339–354. doi:10.1177/1534650120940014
17. **Тычков А. Ю., Чернышов Д. С., Бофанова Н. С., Алимурадов А. К., Овчинников Д. Л., Сотников А. М.** Применение VR для контроля и коррекции фобических тревожных расстройств, *Измерение. Мониторинг. Управление. Контроль*, 2021, № 4, с. 84–92. doi:10.21685/2307-5538-2021-4-10
18. **Somrak A., Pogacnik M., Guna J.** Suitability and comparison of questionnaires assessing virtual reality-induced symptoms and effects and user experience in virtual environments. *Sensors*, 2021, vol. 21, iss. 4, pp. 1185. doi:10.3390/s21041185
19. **Меньшикова Г. Я., Ковалев А. И.** Векция в виртуальных средах: психологические и психофизиологические механизмы формирования. *Национальный психологический журнал*, 2015, № 4, с. 91–104. doi:10.11621/npj.2015.0409
20. **Yang J. H., Ryu J. J., Nam E., Lee H.-S., Lee J. K.** Effects of preoperative virtual reality magnetic resonance imaging on preoperative anxiety in patients undergoing arthroscopic knee surgery: a randomized controlled study. *Arthrosc. J. Arthrosc. Relat. Surg.*, 2019, vol. 35, iss. 8, pp. 2394–2399. doi:10.1016/j.arthro.2019.02.037
21. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (дата обращения: 11.03.2022).
22. *Integrated development environment for R*. Studio, PBC, Boston, MA. http://www.rstudio.com/ (дата обращения: 11.03.2022).
23. *Extension to 'ggplot2'. R package version 2.1.2*. https://CRAN.R-project.org/package=GGally (дата обращения: 18.04.2022).
24. *Generate alluvial plots with a single line of code. R package version 0.3.0*. https://CRAN.R-project.org/package=easyalluvial (дата обращения: 18.04.2022).

UDC 004.946
doi:10.31799/1684-8853-2022-4-58-67

Search for EEG signal patterns in simulating phobic anxiety disorder situations in a VR environment

A. Yu. Tychkov[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-2354-2895, tychkov-a@mail.ru
D. S. Chernyhov[a], Programmer Engineer, orcid.org/0000-0002-1604-7001
P. P. Churakov[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-5918-732X
Z. M. Yuldashev[b], Dr. Sc., Tech., Professor, orcid.org/0000-0003-1075-3420
N. S. Bofanova[a], PhD, Med., Associate Professor, orcid.org/0000-0002-5455-4987
A. K. Alimuradov[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-5133-2713
V. N. Gorbunov[a], PhD, Econ., Associate Professor, orcid.org/0000-0003-0195-0772
R. V. Zolotarev[a], Programmer Engineer, orcid.org/0000-0002-5614-6557
M. S. Nikitin[a], Programmer Engineer, orcid.org/0000-0002-5708-3391
[a]Penza State University, 40, Krasnaya St., 440000, Penza, Russian Federation
[b]Saint-Petersburg Electrotechnical University «LETI», 5, Prof. Popov St., 197376, Saint-Petersburg, Russian Federation

Introduction: Effective and prompt formulation of diagnostic conclusions about the presence of anxiety-phobic disorders requires the improvement of existing and the development of new methods for diagnosing and treating patients, including the use of virtual reality technology. **Purpose:** To analyze a reaction of an individual to a stimulus that triggers a fear response to virtual reality scenes (height exposure). To identify electroencephalographic (EEG) signal markers related to the level of anxiety and virtual reality environment susceptibility of an individual. **Methods:** A group of nine conditionally healthy males aged 23 to 26 years old who reported neither history of somatic symptoms nor organic brain disorders was formed to conduct the research. The immersion into virtual reality was accompanied by the registration of EEG signals and subsequent completion of a self-assessment questionnaire by the subjects. **Results:** The state of rest (a reference value) and the state of high emotional stress experience (at the height of a skyscraper) in the virtual reality environment were compared. The results obtained allow to make a conclusion that the simulated situation of being at a height causes a decrease in the indices of alpha, theta, beta rhythms, and an increase in the delta rhythm index of the EEG signal relative to the state of rest in various subjects, regardless of the intensity of fear manifestation. **Practical relevance:** The conducted research is among the pioneering studies in assessing the effect of virtual reality technologies on human phobic anxiety state. Some objective electrophysiological markers related to the level of anxiety were determined to confirm the presence of patterns in the functional state of the cerebral cortex with a sense of anxiety in individuals immersed in a virtual reality environment.

Keywords — virtual reality, phobic anxiety disorders, electroencephalographic signal.

For citation: Tychkov A. Yu., Chernyhov D. S., Churakov P. P., Yuldashev Z. M., Bofanova N. S., Alimuradov A. K., Gorbunov V. N., Zolotarev R. V., Nikitin M. S. Search for EEG signal patterns in simulating phobic anxiety disorder situations in a VR environment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 58–67 (In Russian). doi:10.31799/1684-8853-2022-4-58-67

Financial support

The work was supported by a grant from the Fund for the Promotion of the Development of Small Businesses in the Sphere of Science and Technology, № 0015902.

References

1. Karpov O. E., Daminov V. D., Novak E. V., Mychametova D. A., Slepnyova N. I. Virtual reality technologies in medical rehabilitation as an example of modern health informatization. *Bulletin of Pirogov National Medical and Surgical Center*, 2020, vol.15, no. 1, pp. 89–98.
2. Pereira J. S., Faêda L. M., Coelho A. M. Evolution of VRET to assist in the treatment of phobias: A systematic review. *22nd Symp. on Virtual and Augmented Reality*, 2020, pp. 386–390. doi:10.1109/SVR51698.2020.00064
3. Li Y., Chiu P., Yeh S., Zhou C., Chen J. Effects of virtual reality and augmented reality on induced anxiety. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, vol. 26, no. 7, pp. 1–8. doi:10.1109/ES.2017.28
4. Krohn S., Tromp J., Quinque E. M., Belger J., Klotzsche F., Gaebler M., Thöne-Otto A., Finke C. Multidimensional assessment of virtual reality applications in clinical neuropsychology: The "VR-check" protocol. *Intern. Conf. on Virtual Rehabilitation*, 2019, 12 p. doi:10.1109/ICVR46560.2019.8994590
5. Tychkov A. Yu., Chernyshov D. S., Bofanova N. S., Alimuradov A. K., Ovchinnikov D. L., Sotnikov A. M. Virtual reality implementation for assessment and treatment of phobic anxiety disorders. *5th Scientific School Dynamics of Complex Networks and their Applications*, 2021, pp. 202–205. doi:10.1109/DCNA53427.2021.9587127
6. Bofanova N. S., Bulanov A. A., Yavorsky A. S., Alyokhina E. V. Virtual reality technology as a modern direction in the rehabilitation of patients with phantom pain. *Russian Journal of Pain*, 2021, no. 19(2), pp. 33–37 (In Russian). doi:10.17116/pain20211902133
7. Karavaeva T. A., Vasilyeva A. V., Mizinova E. V., Poltorak S. V., Lyashkovskaya S. V. Clinical recommendations Anxiety-phobic disorders F40. Moscow, Russian Society of Psychiatrists, 2015. 50 p. (In Russian).
8. Botella C., Fernandez-Alvarez J., Guillén V., García-Palacios A., Baños R. Recent progress in virtual reality exposure therapy for phobias: A systematic review. *Current Psychiatry Reports*, 2017, vol. 19, iss. 7, 13 p. doi:10.1007/s11920-017-0788-4
9. Indovina I., Conti A., Lacquaniti F., Staab J. P., Passamonti L., Toschi N. Lower functional connectivity in vestibular-limbic networks in individuals with subclinical agoraphobia. *Frontiers in Neurology*, 2019, vol. 10, 13 p. doi:10.3389/fneur.2019.00874
10. Hasani J., Rasti A. The comparison of implicit and explicit memory bias to negative information processing in adolescents with high and low trait anxiety. *Journal of Psychology*, 2010, vol. 4, pp. 11–22.
11. Gaur S., Panjwani U., Kumar B. EEG brain wave dynamics: A systematic review and meta analysis on effect of yoga on mind relaxation. *J Biomed Res Environ Sci*, 2020, vol. 1, iss. 7, pp. 353–362. doi:10.37871/jbres1166
12. Freeman D., Reeve S., Robinson A., Ethlers A., Clark D., Spanlang B., Slater M. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 2017, vol. 47, iss. 14, pp. 2393–2400. doi:10.1017/S003329171700040X
13. Jang D. P., Ku J. H., Choi Y. H., Wiederhold B. K., Nam S. W., Kim In Y., Kim Sun I. The development of virtual reality therapy (VRT) system for the treatment of acrophobia and

therapeutic case. *IEEE Trans Inf Technol Biomed*, 2002, pp. 213–220. doi:10.1109/titb.2002.802374
14. Jadhav L., Parekh K., Gupta V., Chandak S. Using virtual reality for therapeutic treatment of phobia. *IEEE Intern. Conf. on Convergence to Digital World — Quo Vadis*, 2020, 8 p. doi:10.1109/ICCDW45521.2020.9318727
15. Bofanova N. S., Petrova E. V., Kalistratov V. B., Nesterenko E. N., Chizh D. I. Using virtual reality for pain management in children. *Ul'yanovskiy mediko-biologicheskiy zhurnal*, 2020, no. 4, pp. 19–29 (In Russian). doi:10.34014/2227-1848-2020-4-19-29
16. Malbos E., Burgess G., Lançon C. Virtual reality and fear of shark attack: A case study for the treatment of squalophobia. *Clinical Case Studies*, vol. 19, pp. 339–354. doi:10.1177/1534650120940014
17. Tychkov A. Yu., Chernyshov D. S., Bofanova N. S., Alimuradov A. K., Ovchinnikov D. L., Sotnikov A. M. VR implementation for control and correct phobic anxiety disorders. *Measurements. Monitoring. Management. Control*, 2021, no. 4, pp. 84–92 (In Russian). doi:10.21685/2307-5538-2021-4-10
18. Somrak A., Pogacnik M., Guna J. Suitability and comparison of questionnaires assessing virtual reality-induced symptoms and effects and user experience in virtual environments. *Sensors*, 2021, vol. 21, iss. 4, pp. 1185. doi:10.3390/s21041185
19. Menshikova G. Y., Kovalev A. I. Vection in virtual environments: psychologicall and psychophysiological mechanisms. *National Psychological Journal*, 2015, no. 4, pp. 91–104 (In Russian). doi:10.11621/npj.2015.0409
20. Yang J. H., Ryu J. J., Nam E., Lee H.-S., Lee J. K. Effects of preoperative virtual reality magnetic resonance imaging on preoperative anxiety in patients undergoing arthroscopic knee surgery: a randomized controlled study. *Arthrosc. J. Arthrosc. Relat. Surg.*, 2019, vol. 35, iss. 8, pp. 2394–2399. doi:10.1016/j.arthro.2019.02.037
21. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/ (accessed 11 March 2022).
22. *Integrated development environment for R*. Studio, PBC, Boston, MA. Available at: http://www.rstudio.com/ (accessed 11 March 2022).
23. *Extension to 'ggplot2'. R package version 2.1.2*. Available at: https://CRAN.R-project.org/package=GGally (accessed 18 April 2022).
24. *Generate alluvial plots with a single line of code*. R package version 0.3.0. Available at: https://CRAN.R-project.org/package=easyalluvial (accessed 18 April 2022).

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая Scopus и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, что снижает рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12 языках, включая русский (чтобы выбрать язык, кликните на зеленое поле вверху справа на стартовой странице): https://orcid.org

СВЕДЕНИЯ ОБ АВТОРАХ

АЛИМУРАДОВ Алан Казанферович

Доцент кафедры радиотехники и радиоэлектронных систем Пензенского государственного университета.
В 2007 году окончил Пензенский государственный университет по специальности «Радиотехника».
В 2015 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 200 научных публикаций и семи патентов на изобретения.
Область научных интересов — разработка и практическое применение способов цифровой обработки речевых сигналов, распознавания речи, голосового управления, голосовой аутентификации, преобразования Гильберта — Хуанга, вейвлет-преобразования и преобразования Фурье.
Эл. адрес: alansapfir@yandex.ru

ВИШНЕВСКИЙ Владимир Миронович

Профессор, заведующий лабораторией Института проблем управления им. В. А. Трапезникова РАН, Москва, заслуженный деятель науки Российской Федерации.
В 1971 году окончил факультет прикладной математики Московского института электронного машиностроения по специальности «Инженер-математик».
В 1989 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 300 научных публикаций, десяти монографий и восьми патентов на изобретения и программы ЭВМ.
Область научных интересов — теория и практика построения инфокоммуникационных систем и сетей, прикладная теория вероятностей, теоретическая информатика и др.
Эл. адрес: vishn@ipu.ru

ГОРБУНОВ Валерий Николаевич

Директор центра трансфера технологий Научно-исследовательского института фундаментальных и прикладных исследований Пензенского государственного университета.
В 2003 году окончил Пензенский государственный университет архитектуры и строительства по специальности «Экономика и управление на предприятии».
В 2006 году защитил диссертацию на соискание ученой степени кандидата экономических наук.
Является автором 100 научных публикаций.
Область научных интересов — цифровая обработка сигналов, статистическая обработка и анализ клинических и экспериментальных данных.
Эл. адрес: ctt@pnzgu.ru

БОФАНОВА Наталия Сергеевна

Доцент кафедры неврологии, нейрохирургии и психиатрии Пензенского государственного университета.
В 2010 году окончила Пензенский государственный университет по специальности «Лечебное дело».
В 2016 году защитила диссертацию на соискание ученой степени кандидата медицинских наук.
Является автором 50 научных публикаций.
Область научных интересов — медицинская реабилитация хронических болевых синдромов, тревожно-фобических расстройств с применением технологии виртуальной реальности.
Эл. адрес: neurology-pgu@mail.ru

ГИЗАТУЛИН Азат Ринатович

Доцент кафедры телекоммуникационных систем Уфимского государственного авиационного технического университета.
В 2016 году окончил Уфимский авиационный институт по специальности «Инфокоммуникационные технологии и системы связи».
В 2020 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 60 научных публикаций и одного авторского свидетельства на изобретения.
Область научных интересов — волоконная оптика, телекоммуникационные системы.
Эл. адрес: azat_poincare@mail.ru

ГОРБУНОВА Анастасия Владимировна

Старший научный сотрудник Института проблем управления им. В. А. Трапезникова РАН, Москва.
В 2010 году окончила магистратуру Российского университета дружбы народов по специальности «Прикладная математика и информатика».
В 2017 году защитила диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором 39 научных публикаций.
Область научных интересов — математическая теория телетрафика, теория массового обслуживания, прикладная теория вероятностей, имитационное моделирование и др.
Эл. адрес: avgorbunova@list.ru

ДИНЬ
Зуи Чыонг

Преподаватель факультета информационной безопасности Технологического института почты и телекоммуникаций, Ханой, Вьетнам.
В 2016 году окончил магистратуру по специальности «Инфокоммуникационные технологии и системы связи».
В 2020 году защитил диссертацию на соискание ученой степени кандидата технических наук в Санкт-Петербургском государственном университете телекоммуникаций им. проф. М. А. Бонч-Бруевича.
Область научных интересов — FANET, информационная безопасность.
Эл. адрес: duydt@ptit.edu.vn

ЕРАШОВ
Алексей
Алексеевич

Младший научный сотрудник лаборатории технологий больших данных социокиберфизических систем Санкт-Петербургского Федерального исследовательского центра РАН, Санкт-Петербургский институт информатики и автоматизации РАН.
В 2021 году окончил магистратуру Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Управление в технических системах».
Является автором 18 научных публикаций.
Область научных интересов — методы и алгоритмы управления робототехническими средствами с применением технического зрения и машинного обучения.
Эл. адрес: erashov.a@iias.spb.su

КРЕСТОВНИКОВ
Константин
Дмитриевич

Младший научный сотрудник лаборатории автономных робототехнических систем Санкт-Петербургского Федерального исследовательского центра РАН, Санкт-Петербургский институт информатики и автоматизации РАН.
В 2019 году окончил магистратуру Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Мехатроника и робототехника».
Является автором 38 научных публикаций и двух патентов на изобретения.
Область научных интересов — модели и алгоритмы распределения энергетических ресурсов робототехнических средств путем двунаправленной беспроводной передачи энергии, мехатроника.
Эл. адрес:
krestovnokov.@iias.spb.su

ДО
Фук Хао

Преподаватель факультета информационных технологий Данангского архитектурного университета, Дананг, Вьетнам.
В 2017 году окончил магистратуру Данангского университета науки и технологий, Вьетнам, по специальности «Компьютерные науки».
Является автором шести научных публикаций.
Область научных интересов — искусственный интеллект, блокчейн, информационная безопасность.
Эл. адрес: haodp@dau.edu.vn

ЗОЛОТАРЕВ
Руслан
Валерьевич

Инженер кафедры радиотехники и радиоэлектронных систем Пензенского государственного университета.
Является автором 15 научных публикаций.
Область научных интересов — разработка AR-, VR-технологий для практического использования в медицине (когнитивные технологии), образовательных программах и научно-исследовательской деятельности.
Эл. адрес:
leshiy021178@gmail.com

КУЗЬМИН
Владимир
Никифорович

Профессор, ведущий сотрудник управления военно-исследовательского (научного) института Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург, заслуженный деятель науки Российской Федерации.
В 1976 году окончил Военный инженерный Краснознаменный институт им. А. Ф. Можайского по специальности «Автоматизированные системы управления».
В 1996 году защитил диссертацию на соискание ученой степени доктора военных наук.
Является автором 200 научных публикаций.
Область научных интересов — теория управления, автоматизированные системы управления.
Эл. адрес: vka@mil.ru

СВЕДЕНИЯ ОБ АВТОРАХ

ЛЕ
Дык Чан

Лектор факультета информационных технологий Института науки и техники Данангского университета, Вьетнам.
В 2014 году окончил Санкт-Петербургский государственный университет телекоммуникаций им. профессора М. А. Бонч-Бруевича по специальности «Многоканальные телекоммуникационные системы».
В 2018 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 30 научных публикаций.
Область научных интересов — беспроводная сеть, VANET, FANET, SDN, анализ вредоносных программ.
Эл. адрес: letranduc@dut.udn.vn

МЕШКОВ
Иван
Константинович

Доцент кафедры телекоммуникационных систем Уфимского государственного авиационного технического университета.
В 2007 году окончил Уфимский авиационный институт по специальности «Радиосвязь, радиовещание и телевидение».
В 2010 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 80 научных публикаций и пяти авторских свидетельств на изобретения.
Область научных интересов — обработка сигналов, спутниковые данные, обработка цифровых сигналов.
Эл. адрес: mik.ivan@bk.ru

СУЛТАНОВ
Альберт
Ханович

Профессор, заведующий кафедрой телекоммуникационных систем Уфимского государственного авиационного технического университета, заслуженный деятель науки Республики Башкортостан, заслуженный работник высшей школы РФ.
В 1973 году окончил Новосибирский электротехнический институт связи по специальности «Многоканальная электросвязь». В 1997 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 200 научных публикаций и 16 патентов на изобретения.
Область научных интересов — полупроводниковая и антенная техника, оптика, связь, обработка изображений.
Эл. адрес: tks@ugatu.ac.ru

МЕНИСОВ
Артем
Бакытжанович

Докторант кафедры систем сбора и обработки информации Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 2013 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Информационные системы и технологии».
В 2018 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 51 научной публикации.
Область научных интересов — искусственный интеллект, машинное обучение, нейронные сети, информационная безопасность.
Эл. адрес: men.arty@yandex.ru

НИКИТИН
Михаил
Сергеевич

Студент, лаборант кафедры радиотехники и радиоэлектронных систем Пензенского государственного университета.
Является автором 10 научных публикаций.
Область научных интересов — статистическая обработка и анализ экспериментальных данных.
Эл. адрес:
deniska_1980_13@mail.ru

ТИМОФЕЕВ
Александр
Леонидович

Доцент кафедры электроники и биомедицинских технологий Уфимского государственного авиационного технического университета.
В 1977 году окончил Уфимский авиационный институт по специальности «Информационно-измерительная техника».
В 1984 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 52 научных публикаций и 13 авторских свидетельств на изобретения.
Область научных интересов — помехоустойчивое кодирование, связь, обработка изображений, цифровая голография.
Эл. адрес: a_l_t@inbox.ru

ТЫЧКОВ Александр Юрьевич

Профессор кафедры радиотехники и радиоэлектронных систем Пензенского государственного университета.
В 2009 году окончил Пензенский государственный университет по специальности «Инженерное дело в медико-биологической практике».
В 2019 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 250 научных публикаций и 12 патентов на изобретения.
Область научных интересов — разработка и практическое применение способов цифровой обработки медицинских сигналов и инструментов на их основе, в том числе с применением технологии виртуальной реальности.
Эл. адрес: tychkov-a@mail.ru

ЧЕРНЫШОВ Денис Сергеевич

Студент, лаборант кафедры радиотехники и радиоэлектронных систем Пензенского государственного университета.
Является автором 15 научных публикаций.
Область научных интересов — цифровая обработка сигналов, клинические исследования в области виртуальной реальности.
Эл. адрес: deniska_1980_13@mail.ru

ЮЛДАШЕВ Зафар Мухамедович

Профессор, заведующий кафедрой биотехнических систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ», почетный работник высшего профессионального образования.
В 1978 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Электронно-медицинская аппаратура».
В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 200 научных публикаций и 17 патентов на изобретения.
Область научных интересов — биотехнические системы медицинского назначения, медицинское приборостроение, системы удаленного мониторинга состояния здоровья.
Эл. адрес: zmyuldashev@etu.ru

ФАМ Мау Хоа

Инженер отдела исследований и разработок, Viettel Business Solutions Корпорация, Дананг, Вьетнам.
В 2020 году окончил Университет науки и техники Данангского университета по специальности «Информационные технологии».
Область научных интересов — беспроводная сеть, машинное обучение, блокчейн, вредоносное программное обеспечение.
Эл. адрес: mauph@viettel.com.vn

ЧУРАКОВ Петр Павлович

Профессор кафедры информационно-измерительной техники и метрологии Пензенского государственного университета.
В 1968 году окончил Куйбышевский политехнический институт по специальности «Электронные устройства управления».
В 1998 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 300 научных публикаций и 25 патентов на изобретения.
Область научных интересов — информационные системы измерения, управления и контроля с использованием параметрических и генераторных первичных преобразователей различных физических величин, устройства генерации, формирования и обработки широкополосных сигналов.
Эл. адрес: churakov-pp@mail.ru

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Текст рукописи должен быть оригинальным, а цитирование и самоцитирование корректно оформлено.

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисуночные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени.

Простые **формулы** набирайте в Word, сложные с помощью редактора Mathtype или Equation. Для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта в Mathtype никогда не пользуйтесь вкладкой Other, Smaller, Larger, используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; пробелы в формуле ставьте только после запятых при перечислении с помощью Ctrl+Shift+Space (пробел); не отделяйте пробелами знаки: + = – ×, а также пространство внутри скобок; для выделения греческих символов в Mathtype полужирным начертанием используйте Style → Other → bold.

Для набора формул в Word никогда не используйте вкладки: «Уравнение», «Конструктор», «Формула» (на верхней панели: «Вставка» — «Уравнение»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Подробнее см. pdf-файл «Правила подготовки рукописей» (стр. 11) на сайте https://guap.ru/ric

Иллюстрации:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Adobe Illustrator (*.ai); Coreldraw (*.cdr, версия не выше 15); Excel (*.xls); Word (*.docx); AutoCad, Matlab (экспорт в PDF, EPS, SVG, WMF, EMF); Компас (экспорт в PDF), веб-портал DRAW.IO (экспорт в PDF);

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисуночных подписей и названий таблиц на русском и английском языках обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png, *.jpg с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц, doi;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц, doi;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (http://i-us.ru/paper-rules): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Руководство для авторов».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: ius.spb@gmail.com
Сайт: www.i-us.ru