16+

ISSN 2074-1707

АСТРАХАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2022
№ 1 (57)

Журнал включен в перечень рецензируемых научных изданий, рекомендованных ВАК России для публикации основных научных результатов диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук по следующим научным специальностям.

Группа специальностей 1.2 «Компьютерные науки и информатика»:

1.2.2 – Математическое моделирование, численные методы и комплексы программ (технические науки).

Группа специальностей 2.2 «Электроника, фотоника, приборостроение и связь»:

2.2.4 – Приборы и методы измерения (по видам измерений) (технические науки);

2.2.11 – Информационно-измерительные и управляющие системы (технические науки);

2.2.12 – Приборы, системы и изделия медицинского назначения (технические науки).

Группа специальностей 2.3 «Информационные технологии и телекоммуникации»:

2.3.1 – Системный анализ, управление и обработка информации (технические науки);

2.3.4 – Управление в организационных системах (технические науки);

2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки);

2.3.6 – Методы и системы защиты информации, информационная безопасность (технические науки).

Журнал входит в базу данных Ulrich’s Periodicals Directory.

Астрахань
Издательский дом «Астраханский университет»
2022

ББК 30
П75

Рекомендовано к печати редакционно-издательским советом
Астраханского государственного университета

ПРИКАСПИЙСКИЙ ЖУРНАЛ: управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2022
№ 1 (57)

Журнал выходит 4 раза в год
Все материалы, поступающие в редколлегию журнала,
проходят независимое рецензирование

ISSN 2074-1707

ASTRAKHAN STATE UNIVERSITY

PRIKASPIYSKIY ZHURNAL:
Upravlenie i Vysokie Tekhnologii

CASPIAN JOURNAL:
Control and High Technologies

A SCIENTIFIC AND TECHNICAL JOURNAL

2022
No. 1 (57)

The journal is included in the list of the reviewed scientific journals recommended by VAK of Russia for the publication of the main scientific results of theses for the candidate of science degree, for the doctor of science degree on the following scientific specialties.

Group of specialties 1.2 "Computer science and informatics":

1.2.2 – Mathematical modelling, numerical methods and complexes of programms (technical sciences).

Group of specialties 2.2 "Electronics, photonics, instrument engineering and communication":

2.2.4 – Instruments and methods of measurement (by type of measurement) (technical sciences);

2.2.11 – Information-measuring and control systems (technical sciences);

2.2.12 – Medical devices, systems and products (technical sciences).

Group of specialties 2.3 "Information technologies and telecommunications":

2.3.1 – System analysis, information control and processing (technical sciences);

2.3.4 – Management in organizational systems (technical sciences);

2.3.5 – Mathematical software and software for computing systems, complexes and computer networks (technical sciences);

2.3.6 – Information security methods and systems, information security (technical sciences).

The journal is included into the database Ulrich's Periodicals Directory.

Astrakhan
Publishing House "Astrakhan University"
2022

Recommended by the Editorial and Publishing Board
of Astrakhan State University

CASPIAN JOURNAL:
Control and High Technologies

A SCIENTIFIC AND TECHNICAL JOURNAL

2022
No. 1 (57)

The journal is published four times a year
All materials that come to the Editorial Board of the journal
are subject to independent peer-review

СОДЕРЖАНИЕ

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

УПРАВЛЕНИЕ В ОРГАНИЗАЦИОННЫХ СИСТЕМАХ

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ МАШИН, КОМПЛЕКСОВ И КОМПЬЮТЕРНЫХ СЕТЕЙ

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И ЗАЩИТА ИНФОРМАЦИИ

CONTENTS

INFORMATICS, COMPUTER TECHNIQUE AND CONTROL

SYSTEM ANALYSIS, CONTROL AND INFORMATION PROCESSING

MANAGEMENT IN ORGANIZATIONAL SYSTEMS

MATHEMATICAL MODELLING, NUMERICAL METHODS AND PROGRAM SYSTEMS

MATHEMATICAL SOFTWARE AND SOFTWARE FOR COMPUTING MACHINES, COMPLEXES AND COMPUTER NETWORKS

INFORMATION SAFETY AND INFORMATION PROTECTION

INSTRUMENT ENGINEERING, MEASUREMENT SCIENCE, INFORMATION AND MEASURING DEVICES AND SYSTEMS

INFORMATION-MEASURING AND CONTROL SYSTEMS

MEDICAL DEVICES, SYSTEMS AND PRODUCTS

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

DOI 10.54398/2074-1707_2022_1_9
УДК 004.4

ПАРСИНГ ТЕЛЕГРАМ-КАНАЛОВ КАК ЭЛЕМЕНТ СИСТЕМЫ АВТОМАТИЗИРОВАННОГО АНАЛИЗА ИНФОРМАЦИИ, ПОЛУЧЕННОЙ ИЗ СЕТИ ИНТЕРНЕТ

Статья поступила в редакцию 17.01.2022, в окончательном варианте – 14.02.2022.

Карабак Илья Игоревич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
магистрант, ORCID: 0000-0002-5894-9273, e-mail: razvitienonedegradacia@gmail.com
Зорин Кирилл Андреевич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
ассистент кафедры информационной безопасности и цифровых технологий, ORCID: 0000-0003-1614-8168, e-mail: kirocan95@gmail.com
Ажмухамедов Искандар Маратович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
доктор технических наук, декан факультета цифровых технологий и кибербезопасности, профессор кафедры информационной безопасности, ORCID: 0000-0001-9058-123X, e-mail: aim_agtu@mail.ru

В данной статье были описаны специфика работы и протоколы одного из самых популярных мессенджеров – Telegram. Разработанное программное обеспечение позволяет автоматизировать сбор текстовых данных из публичных сообществ мессенджера Telegram. Для реализации сервисного взаимодействия был использован брокер сообщений RabbitMQ. Все полученные сообщения из мессенджера передаются в очередь брокера, к которой при наличии конфигурационных данных может подключиться любой сервис – анализатор и начать обрабатывать полученные данные. Реализованное программное обеспечение актуально для автоматизированной обработки информации, публикуемой в мессенджере.
Ключевые слова: анализ, Telegram, параметры, парсинг, короткие сообщения, сервис

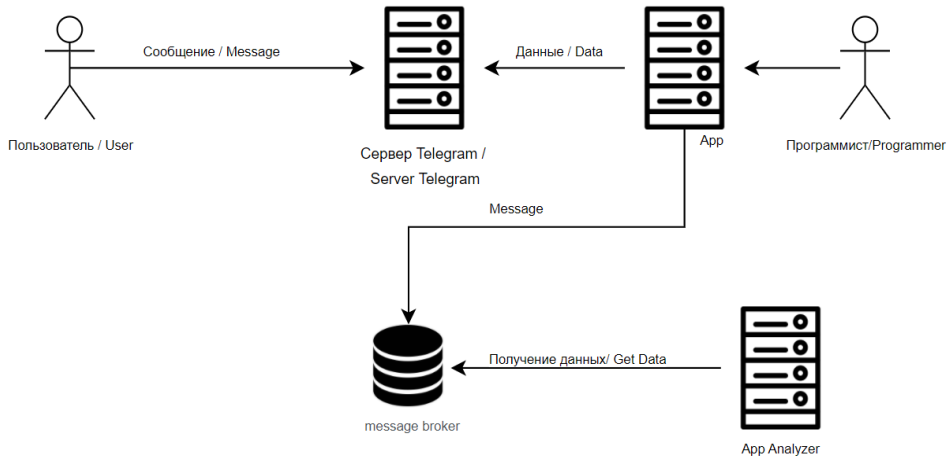
PARSING OF TELEGRAM CHANNELS AS AN ELEMENT OF THE SYSTEM OF AUTOMATED ANALYSIS OF INFORMATION RECEIVED FROM THE INTERNET

The article was received by the editorial board on 17.01.2022, in the final version – 14.02.2022.

Karabak Ilya I., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
undergraduate student, ORCID https://orcid.org/0000-0002-5894-9273, e-mail: razvitienon edegradacia@gmail.com
Zorin Kirill A., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
Assistant of the Department of Information Security and Digital Technologies, e-mail: kirocan95@gmail.com
Azhmukhamedov Iskandar M., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
Doct. Sci. (Engineering), Dean of the Faculty of Digital Technologies and Cybersecurity, Professor of the Department of Information Security, e-mail: aim_agtu@mail.ru

This article described the specifics of the work and described the protocols of one of the most popular instant messengers – Telegram. The developed software allows you to automate the collection of text data from the public communities of the Telegram messenger. To implement the service interaction, a message broker, RabbitMQ, was used. All received messages from the messenger are transferred to the broker's queue, to which, if configuration data is available, any analyzer service can connect and start processing the received data. The implemented software is relevant for automated processing of information published in the messenger
Keywords: analysis, Telegram, parameters, parsing, short messages, service

Graphical annotation (Графическая аннотация)



Введение. На сегодняшний день интернет-пространство стало неотъемлемой частью жизни современного человека. Объемы информации, которую получает человек в современном мире, с каждым днём увеличиваются в геометрической прогрессии. В интернет-среде существуют различные источники информации, такие как онлайн-СМИ, социальные сети, мессенджеры, чаты, форумы. Информация в этих источниках может содержать мнения пользователей; общедоступные данные, размещаемые пользователями в социальных сетях; результаты поисковой выдачи и т. д.

Одним из наиболее распространённых источников распространения информации в интернет-пространстве является мессенджер Telegram, который в последние годы приобрел большую популярность в мире. Только за 4 года, с 2016 по 2020 г., аудитория данного мессенджера увеличилась более чем в 4 раза (рис. 1).

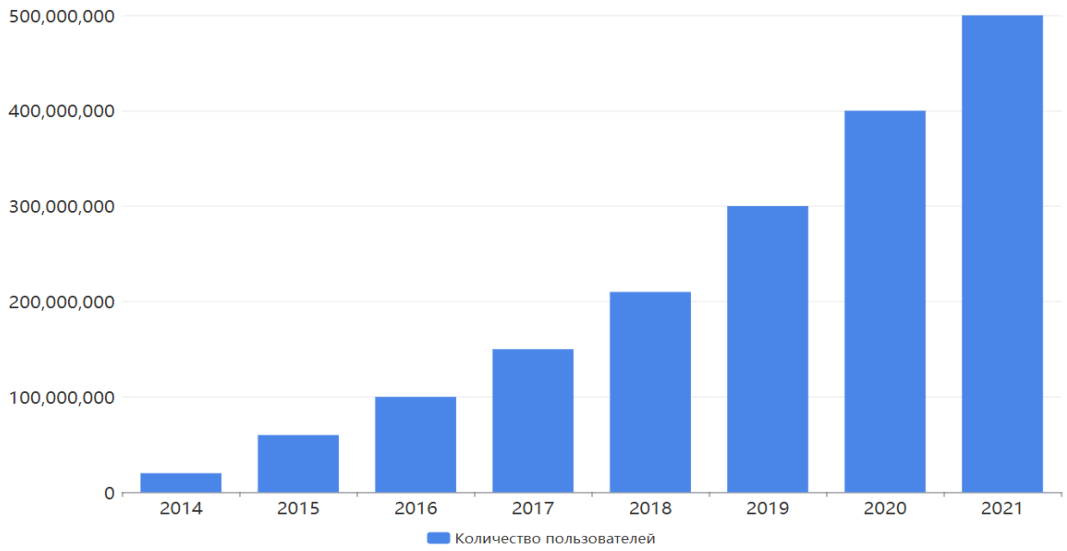


Рисунок 1 – Рост пользователей мессенджера Telegram по данным сайта telegram.org [1]

При этом среди пользователей мессенджера наиболее популярны Telegram-каналы. Существует два вида Telegram-каналов:

- публичные – общедоступный вид сообществ – такие каналы доступны любому пользователю мессенджера для чтения и подписки. Это самый распространенный формат сообществ в Telegram;
- частные – закрытый вид сообществ, доступ в которые можно получить только по ссылке, известной только администратору канала.

Пользователей Telegram привлекает его анонимность, безопасность, понятный интерфейс,

благодаря которым любой желающий может делиться с другими пользователями тем или иным контентом. Как следствие, это приводит к увеличению количества источников информации в мессенджере. Ввиду роста объема информации возникает потребность автоматизации ее анализа и систематизации, что может представлять интерес для различных организаций, например, для бизнеса или правоохранительных органов. К примеру, компании могут проводить анализ лояльности клиентов к их продукту или услуге, основываясь на комментариях пользователей в мессенджере, а правоохранительные органы – выявлять неправомерную информацию. При этом ручной сбор информации является весьма трудоемкой задачей ввиду больших объемов данных. Одним из возможных методов решения этой проблемы является автоматизация процесса сбора информации по заданным параметрам – парсинг.

В статье Т.Э. Вильданова, Н.С. Иванова «Анализ инструментов парсинга и веб-скрейпинга в рамках разработки арбитражной инвестиционной стратегии на рынке спортивных ставок» [2] приводится исследование популярных инструментов для парсинга публичных сайтов, а также выявлены наиболее быстрые инструменты и методы для быстрого извлечения и обработки информации в больших количествах. В статье Д.З. Цхошвили, Н.А. Иванова «Примеры использования технологии парсинга» [3] приводятся примеры использования технологии анализа данных – парсинга, которые представлены в современных научных работах. Однако методы и инструменты, описанные в данных статьях, не позволяют производить парсинг публичных Telegram-каналов, так как, в отличие от сайтов, мессенджер Telegram имеет свою сложную структуру хранения сообщений, а также использует отдельный протокол для работы с сообщениями. Таким образом, задача автоматизированного сбора информации, публикуемой в общедоступных группах, созданных в мессенджере Telegram, является не до конца решенной.

Постановка задачи. Исходя из этого, целью данной работы стало описание структуры разработанного программного обеспечения, которое позволит осуществлять сбор информации, публикуемой в общедоступных группах, созданных в мессенджере Telegram.

Программное обеспечение будет являться компонентом общей системы анализа текстовой информации, полученной из различных источников информации интернета. Общая схема взаимодействия компонентов такой системы, построенная в стандарте UML [2], представлена на рисунке 2.

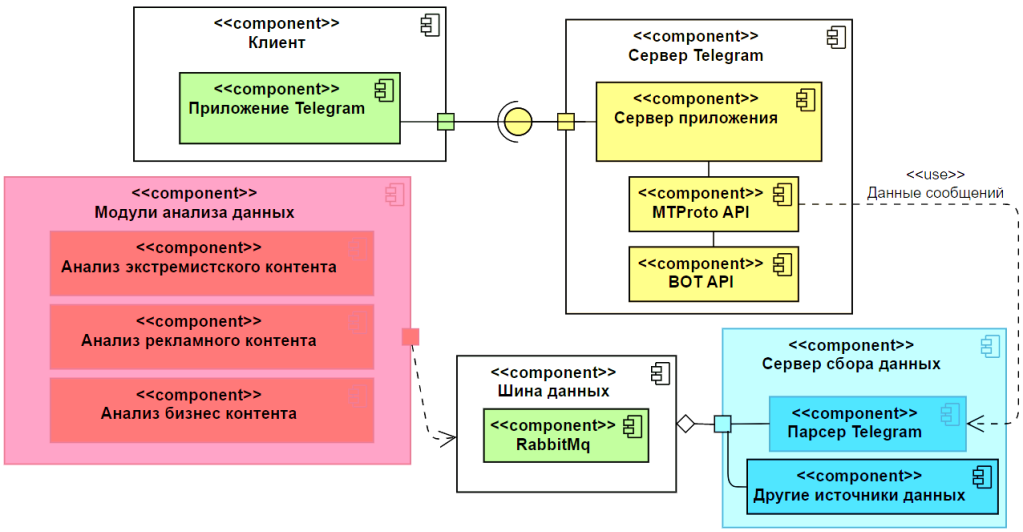


Рисунок 2 – Диаграмма развертывания системы анализа текстовой информации, полученной из различных источников интернета

Прежде чем начать разработку ПО, выполняющего функцию парсинга, требуется рассмотреть специфику мессенджера Telegram.

Специфика мессенджера Telegram. Контент, обрабатывающийся в мессенджере, можно разделить на 3 группы:

- текстовый контент (сообщения, комментарии к сообщениям);
- медиаконтент (видео, аудио);
- прочие файлы (документы, архивы, файлы приложений).

В мессенджере Telegram существует два вида интернет-сообществ: чаты и каналы. Telegram-

канал представляет собой чат с односторонней связью, в котором только администратор канала может оставлять сообщения. Также, по желанию, администратор канала может предоставить доступ подписчика канала комментировать его публикации. Telegram-чат представляет из себя площадку, где все участники могут принимать участие в общение. Возможность комментировать сообщения участников отсутствует.

Для решения задачи автоматизации сбора данных, представляющих собой текстовую информацию, были проанализированы способы и методы, позволяющие взаимодействовать с программным обеспечением мессенджера Telegram. В частности, была исследована документация API-набора функций, методов, которые позволяют взаимодействовать с программным обеспечением Telegram строго в рамках политики компании [3]. Это позволило выбрать возможные пути доступа к сервисам мессенджера, которые могут быть использованы при разработке программного обеспечения для парсинга. Для того чтобы взаимодействовать с API, требуется заполнить специальную форму на сайте https://my.telegram.org/, с помощью которой создаётся собственное приложение Telegram API (рис. 3) [4].

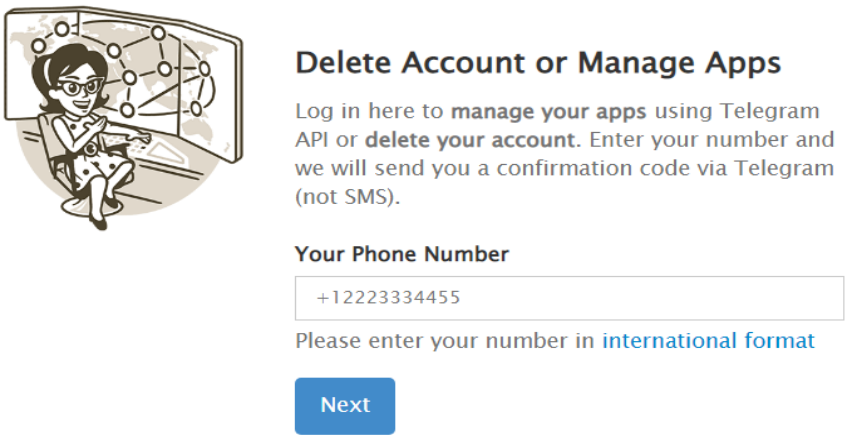


Рисунок 3 – Заполнение формы Telegram API

Приложение хранит уникальные значения, идентифицирующие пользователя в системе Telegram API, – «api_id» и «api_hash». Значение «api_id» представляет собой последовательность цифр, а «api_hash» – буквенно-цифровую последовательность. Оба значения автоматически генерируются сервером Telegram на бессрочный период.

Идентификаторы API являются токенами приложения. Это означает, что при использовании приложения пользователю не нужно вводить собственные идентификаторы API, достаточно идентификатора разработчика.

Для взаимодействия с сервисами мессенджера пользователю предоставляется два вида API:

- MTProto API;
- Bot API.

Bot API был создан для пользователей, которые разрабатывают Telegram-боты. Боты – это специальные аккаунты, которые авторизуются с помощью токенов, а не телефонных номеров. Bot API построен на основе MTProto API, но использует промежуточный HTTP-сервер для обмена запросами с серверами Telegram. В качестве формата данных для взаимодействия с API используется формат JSON. Большинство методов MTProto API, созданных для работы с сообщениями, недоступны в Bot API.

Изначально протокол MTProto был создан командой разработчиков мессенджера как криптографический протокол для шифрования сообщений. Протокол основан на алгоритме шифрования AES [5], а также на криптографическом алгоритме Диффи – Хеллмана [6].

MTProto API является основным API Telegram, с помощью которого пользователи могут подключаться напрямую к сервисам Telegram (рис. 4). Данный протокол использует сериализацию данных в бинарный вид.

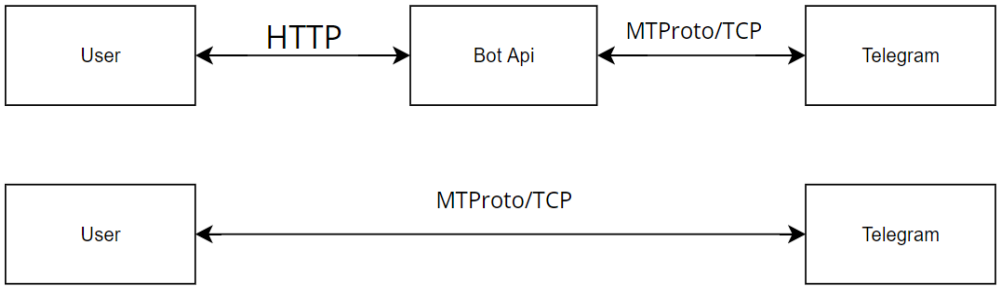


Рисунок 4 – Схема работы протоколов MTProto API и BotAPI

Bot API является упрощенным вариантом MTProto API и используется исключительно для создания ботов, снижая порог входа для обычных пользователей, тогда как MTProto API имеет большой функционал, предлагая пользователю широкие возможности работы с сервисами Telegram. Для задачи парсинга сообщений использовать Bot API невозможно, для этого предлагается использовать протокол MTProto.

Для парсинга сообщений в Telegram-сообществах мессенджер предоставляет различные методы API. Одним из таких является messages.getHistory, которому на вход могут подаваться следующие параметры:

- peer – сущность (идентификатор канала, идентификатор пользователя), от которой разработчик хочет получить сообщения; имеет тип данных “InputPeer”;
- offset_id – идентификатор сообщения; если параметр задан, будут возвращены сообщения, начиная с заданного идентификатора; имеет тип данных целочисленный;
- offset_date – дата отправки сообщения; если параметр задан, будут возвращены только те сообщения, которые были отправлены до указанной даты; имеет тип данных целочисленный;
- add_ofset – данный параметр дополняет параметр offset_id; имеет целочисленный тип данных;
- limit – количество сообщений, которое нужно возвратить; если параметр имеет значение “None”, то метод вовзратит всю историю сообщений; имеет тип данных целочисленный;
- max_id – если данный параметр задан, то метод вернёт все сообщения с меньшими идентификаторами; имеет тип данных целочисленный;
- min_id – если данный параметр задан, то метод вернёт все сообщения с б□о□льшими идентификаторами; имеет тип данных целочисленный;
- hash – данный параметр позволяет хэшировать запрос; имеет целочисленный тип данных.

После того как запрос отправлен, метод messages.getHistory возвращает ответ в виде объекта типа messages.Messages, который содержит в себе список сообщений и информацию об этих сообщениях. Данный объект имеет четыре так называемых конструктора, которые формируют ответы на запрос пользователя в зависимости от значения параметров запроса:

- messages.messages – возвращает полный список сообщений и информацию об этих сообщениях;
- messages.messagesSlice – выдаёт список сообщений и информацию об этих сообщениях;
- messages.channelMessages – возвращает сообщения Telegram-канала;
- messages.messagesNotModified – возвращается в том случае, если сообщения не обнаружены.

В мессенджере Telegram также есть возможность комментировать публикации сообщества, если администратор сообщества даёт доступ к этой функции. Метод messages.getReplies позволяет получить такие комментарии в зависимости от заданных параметров. Основные параметры метода messages.getReplies:

- msg_id – идентификатор сообщения – комментария;
- limit – максимальное количество возвращаемых результатов.

Данный метод также возвращает объект типа messages.Messages в виде списка комментариев в соответствии с заданными параметрами.

Для глобального поиска сообщений мессенджера Telegram существует функция messages.searchGlobal, которая имеет следующие основные параметры:

- folder_id – идентификатор каталога, имеет тип данных flags.0?int;
- q – задаёт значение для глобального поиска, имеет строковый тип данных;
- filter – фильтр сообщений для глобального поиска, тип данных MessagesFilter.

Данный метод также возвращает объект типа messages.Messages.

На основе описанных выше методов было разработано программное обеспечение, которое

позволяет осуществлять сбор текстовой информации, публикуемой в общедоступных группах, созданных в мессенджере Telegram.

Описание ПО. Telegram MTProto API для языка программирования Python реализовано в виде библиотеки Telethon [7]. Данная библиотека содержит в себе множество функций и методов, позволяющих осуществлять работу с серверами мессенджера Telegram.

Для парсинга текстового контента целесообразно использовать функцию iter_messages библиотеки Telethon. С помощью данной функции программа отправляет запрос на сервера Telegram с заданными параметрами. Если параметры верны, программа получает ответ.

Функция iter_messages позволяет задать следующие параметры:

- entity – данный параметр определяет объект, от которого нужно получить сообщение; может быть ссылкой на канал или чат;
- limit – количество сообщений, которое требуется получить; если задано None, то параметр вернёт все сообщения объекта;
- reply_to – комментарии к сообщению, если известен идентификатор сообщения;
- offset_date – возвращает сообщения, предшествующие этой дате.

Взаимодействие программного обеспечения с Telegram API представлено на рисунке 5.

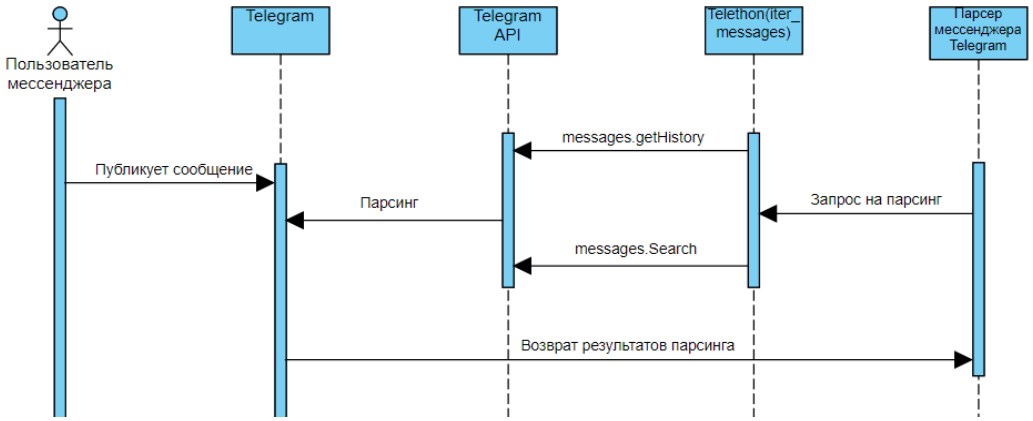


Рисунок 5 – Схема взаимодействия программного обеспечения с Telegram API

Программное обеспечение было реализовано с помощью языка программирования высокого уровня Python версии 3.7 [8]. Для нормального функционирования программе требуется постоянный доступ в интернет. Работа ведётся посредством использования командной строки. Данное программное обеспечение может собирать информацию из интернет-сообществ в мессенджере Telegram, таких как чаты и каналы. Для обоих видов сообществ можно задавать лимит выгружаемых сообщений; период, в течение которого сообщения были опубликованы; имя пользователя, который публиковал сообщения и т.д. После того как результаты парсинга были получены программой-парсером, они выгружаются в брокер сообщений.

Конечная структура ПО выглядит следующим образом (рис. 6).

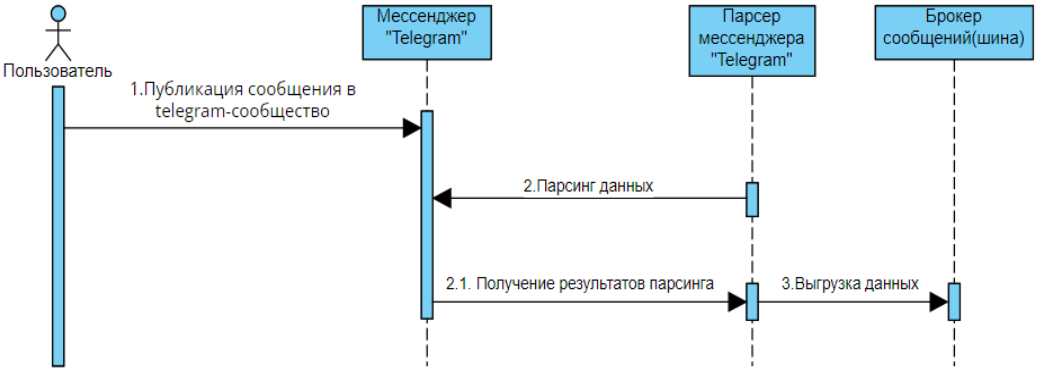


Рисунок 6 – Схема работы парсера

На данной диаграмме отображена схема работы парсера. После того как пользователь мессенджера «Telegram» публикует сообщение в телеграм сообществе, ПО автоматически производит парсинг данных о сообщении, после чего результаты парсинга отправляет в брокер сообщений RabbitMQ, тем самым создавая поток сообщений с возможностью подключения подписчиков.

Пример работы программного обеспечения. Работа с программой осуществляется через командную строку. Для корректного взаимодействия с приложением пользователю необходимо подать на вход программы ссылку на канал или чат из мессенджера Telegram (рис. 7).

Рисунок 7 – Стартовый экран программы

После того как пользователь ввёл ссылку программа начинает парсинг сообщений из канала или чата по ссылке в соответствии с заданными параметрами.

Все данные, полученные в ходе работы парсера, передаются в брокер сообщений RabbitMQ [9]. Брокер хранит в себе данные о публикациях, которые были получены в результате парсинга. Их можно использовать в дальнейшей работе для анализа информации. Для этого сервису-подписчику необходимо подключиться к брокеру, используя строку подключения, введя логин и пароль для доступа, после чего он может использовать передаваемый поток сообщений для проведения анализа и систематизации этих данных. Пример процесса отправки данных в брокер сообщений проиллюстрирован на рисунке 8.

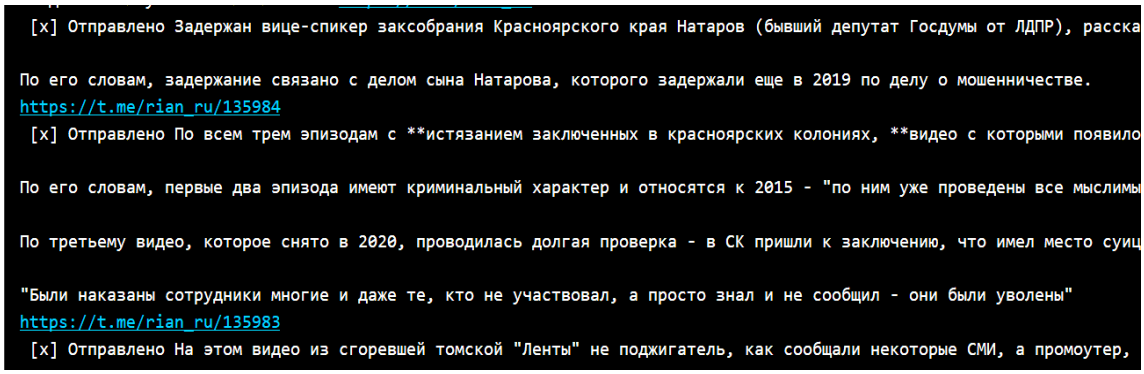


Рисунок 8 – Пример отправки сообщений в брокер сообщений RabbitMQ

Данные, полученные в ходе парсинга, передаются в брокер сообщений в формате данных JSON. Структура сообщения, хранящегося в RabbitMQ, представлена на рисунке 9.

Рисунок 9 – Структура хранения сообщений в RabbitMQ

Сообщение представляет из себя набор пар «ключ – значение», где в качестве значений выступают текстовый контент, полученный в результате парсинга, а также ссылки на полученные данные.

Представленное ПО используется как часть общей системы для анализа текстовой информации. Модуль зарегистрирован как программное обеспечение [12]. Для того чтобы воспользоваться возможностями данного модуля, необходимо получить логин и пароль для доступа к брокеру сообщений и считывать данные в любом режиме. Благодаря данному модулю пользователи смогут автоматизировать процесс сбора данных в различных направлениях, к примеру: сбор комментариев пользователей об услуге, товаре и т. д.

Заключение. При постоянно увеличивающимся потоке информации в интернете проблема автоматизированного сбора публичной информации в телеграм-каналах становится всё более актуальной. В данной статье была описана структура разработанного программного обеспечения, которое позволит осуществлять сбор информации, публикуемой в общедоступных группах, созданных в мессенджере Telegram. Предложенный программный продукт автоматически обрабатывает публикуемые сообщения в сообществах и группах, после чего передает данные в брокер сообщений RabbitMQ. К брокеру сообщений возможно подключить множество программ, которые будут использовать эти данные для различных целей. Основными функциями таких сервисов может быть: обнаружение экстремистских материалов, анализ рекламных блоков, анализ вовлеченности аудитории и т. д. В текущем виде предложенный модуль парсинга телеграм-каналов используется как компонент общей системы для анализа текстовой информации в интернете. Описанный в данной статье программный продукт позволяет экономить ресурсы и средства, а также осуществить в дальнейшем интеграцию модулей для анализа информации, поступающей из разных источников.

Библиографический список

1. 400 миллионов пользователей. – Режим доступа: https://telegram.org/blog/400-million/ru, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 07.12.2021).

2. Вильданов, Т. Э. Анализ инструментов парсинга и веб-скрейпинга в рамках разработки арбитражной инвестиционной стратегии на рынке спортивных ставок / Т. Э. Вильданов, Н. С. Иванов // Скиф. Вопросы студенческой науки. – 2021. – № 5 (57). – С. 23–33.

3. Цхошвили, Д. З. Примеры использования технологии парсинга / Д. З. Цхошвили, Н. А. Иванова // Актуальные вопросы в науке и практике : сборник статей по материалам IV Международной научно-практической конференции : в 5 ч. Самара, 11 декабря 2017 года. – Самара : Общество с ограниченной ответственностью «Дендра», 2017. – С. 135–138.

4. Грейди, Буч. Язык UML. Руководство пользователя = The Unified Modeling Language user guide / Грейди Буч, Джеймс Рамбо, Айвар Джекобсон. – 2-е изд. – Москва ; Санкт-Петербург : ДМК Пресс, Питер, 2004. – 432 с. – ISBN 5-94074-260-2.

5. Telegram APIs. – Режим доступа: https://core.telegram.org/, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 07.12.2021).

6. Delete Account or Manage Apps – Режим доступа: https://my.telegram.org/auth, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 07.12.2021).

7. Баричев, С. Г. 2.4.2. Стандарт AES. Алгоритм Rijdael / С. Г. Баричев, В. В. Гончаров, Р. Е. Серов // Основы современной криптографии. – 3-е изд. – Москва : Диалог-МИФИ, 2011. – С. 30–35. – ISBN 978-5-9912-0182-7.

8. Черёмушкин, А. В. Криптографические протоколы: основные свойства и уязвимости / А. В. Черёмушкин // Прикладная дискретная математика. – 2009. – Вып. 2. – С. 115–150.

9. Telethon’s Documentation. – Режим доступа: https://docs.telethon.dev/en/latest/, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 07.12.2021).

10. Python 3.7.0. – Режим доступа: https://www.python.org/downloads/release/python-370/, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 07.12.2021).

11. RabbitMQ. – Режим доступа: https://www.rabbitmq.com/, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 22.12.2021).

12. Крэг, Ларман. Применение UML 2.0 и шаблонов проектирования / Крэг Ларман. – Вильямс, 2019. – 736 с. – ISBN 978-5-907144-36-1.

13. Ричардсон Крис. Микросервисы. Паттерны разработки и рефакторинга. – Прогресс книга, 2021. – 544 с. – ISBN 978-5-4461-0996-8.

14. Карабак, И. И. Программа для парсинга сообщений в публичных телеграм-каналах : свидетельство о регистрации программы для ЭВМ RU 2021680437, 10.12.2021. – Заявка № 2021669987 от 30.11.2021 / И. И. Карабак, К. А. Зорин, И. М. Ажмухамедов.

References

1. *400 millionov polzovateley* [400 million users]. Available at: https://telegram.org/blog/400-million/ru (accessed 07.12.2021).

2. Vildanov, T. E., Ivanov, N. S. Analiz instrumentov parsinga i veb-skreypinga v ramkahh razrabotki arbitrazhnoy

investitsionnoy strategii na rynke sportivnykh stavok [Analysis of parsing and web-scraping tools as part of the development of an arbitrage investment strategy in the sports betting market]. *Skif. Voprosy studencheskoy nauki* [Scythian Issues of Student Science], 2021, no. 5 (57), pp. 23–33.

3. Tskhoshvili, D. Z., Ivanova, N. A. Primery ispolzovaniya tekhnologii parsinga [Examples of the use of parsing technology]. *Aktualnye voprosy v nauke i praktike : sbornik statey po materialam IV Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Topical issues in science and practice : collection of articles based on the materials of the IV International Scientific and Practical Conference], in 5 parts, Samara, December 11, 2017. Samara, Dendra Limited Liability Company, 2017, pp. 135–138.

4. Grady, Booch, James, Rumbaugh, Ivar, Jacobson. *Yazyk UML. Rukovodstvo polzovatelya* [The Unified Modeling Language. User guide]. 2nd ed. Moscow, St. Petersburg, DMK Press, Peter, 2004. 432 p. ISBN 5-94074-260-2.

5. *Telegram APIs*. Available at: https://core.telegram.org/ (accessed 07.12.2021).

6. *Delete Account or Manage Apps*. Available at: https://my.telegram.org/auth (accessed 07.12.2021).

7. Barichev, S. G., Goncharov, V. V., Serov, R. E. 2.4.2. Standart AES. Algoritm Rijdael [AES standard. Rijdael algorithm]. *Osnovy sovremennoy kriptografii* [Fundamentals of modern cryptography]. 3rd ed. Moscow, Dialogue-MEPhI, 2011, pp. 30–35. ISBN 978-5-9912-0182-7.

8. Cheryomushkin, A. V. Kriptograficheskie protokoly: osnovnye svojstva i uyazvimosti [Cryptographic protocols: basic properties and vulnerabilities]. *Prikladnaya diskretnaya matematika* [Applied Discrete Mathematics], 2009, iss. 2, pp. 115–150.

9. *Telethon's Documentation*. Available at: https://docs.telethon.dev/en/latest/ (accessed 07.12.2021).

10. *Python 3.7.0*. Available at: https://www.python.org/downloads/release/python-370/ (accessed 07.12.2021).

11. *RabbitMQ*. Available at: https://www.rabbitmq.com/ (accessed 12.22.2021).

12. Craig, Larman. *Primenenie UML 2.0 i shablonov proektirovaniya* [Applying UML 2.0 and Design Patterns]. Williams, 2019. 736 p. ISBN 978-5-907144-36-1.

13. Richardson, Chris. *Mikroservisy. Patterny razrabotki i refaktoringa* [Microservices. Patterns of development and refactoring]. Progress book, 2021. 544 p. ISBN 978-5-4461-0996-8.

14. Karabak, I. I., Zorin, K. A., Azhmukhamedov, I.M. *Programma dlya parsinga soobshcheniy v publichnykh telegram-kanalakh : svidetelstvo o registratsii programmy dlya EVM RU 2021680437* [Program for parsing messages in public telegram channels Registration certificate for the computer program RU 2021680437], 10.12.2021. Application No. 2021669987 dated 11.30.2021.

DOI 10.54398/2074-1707_2022_1_17
УДК 004.001

СОЗДАНИЕ КОРПОРАТИВНОЙ ПОЧТОВОЙ СИСТЕМЫ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ УПРАВЛЕНИЯ КОНФИГУРАЦИЯМИ

Статья поступила в редакцию 19.11.2021, в окончательном варианте – 13.01.2022.

Иванова Наталья Александровна, Брянский государственный университет им. академика И.Г. Петровского, 241036, Российская Федерация, г. Брянск, ул. Бежицкая, 14,

кандидат технических наук, доцент, ORCID: 0000-0001-5135-9310, e-mail: ivanova.na@brgu.ru

Кубанских Олеся Владимировна, Брянский государственный университет им. ак. И.Г. Петровского, 241036, Российская Федерация, г. Брянск, ул. Бежицкая, 14,

кандидат физико-математических наук, доцент, ORCID: 0000-0003-0320-1652, e-mail: netbay.ov@brgu.ru

В данной статье рассмотрен процесс развертывания корпоративного почтового сервера посредством системы управления конфигурациями. Проведен анализ наиболее распространенных СМ-систем с учетом используемой архитектуры, наличия web-интерфейса, используемого языка, поддерживаемых платформ, применяемых настроек на контролируемом хосте. В работе были определены технические требования, необходимые для корректного функционирования системы управления конфигурациями и корпоративного почтового сервера. Контроль и управление трафиком почтовой службы вуза, максимальная защита обмена деловой корреспонденции от массовых рассылок и внешних атак, организация комфортного взаимодействия коллег – всё это возможности реализованной корпоративной почтовой системы. Данный способ организации корпоративной почты с использованием системы управления конфигурациями упрощает работу по настройке и развертыванию почтового сервера, оптимизируя работу технического персонала и образовательного учреждения в целом.

Ключевые слова: почтовый сервер, корпоративная почта, система управления конфигурациями, iRedMail, Ansible

CREATING A CORPORATE EMAIL SYSTEM USING A CONFIGURATION MANAGEMENT SYSTEM

The article was received by the editorial board on 19.11.2021, in the final version – 13.01.2022.

Ivanova Nataliya A., Bryansk State University named after Academician I.G. Petrovsky, 14 Bezhitskaya St., Bryansk, 241036, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, ORCID: 0000-0001-5135-9310, e-mail: ivanova.na@brgu.ru

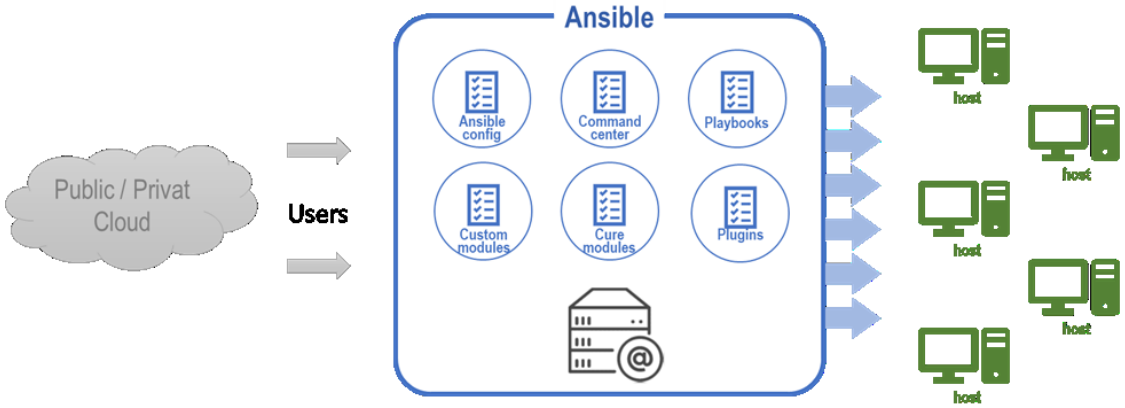
Kubanskikh Olesya V., Bryansk State University named after Academician I.G. Petrovsky, 14 Bezhitskaya St., Bryansk, 241036, Russian Federation,

Cand. Sci. (Physics and Mathematics), Associate Professor, ORCID: 0000-0003-0320-1652, e-mail: netbay.ov@brgu.ru

This article describes the process of deploying a corporate mail server through a configuration management system. The possibilities of the most popular configuration management systems are analyzed in terms of such basic criteria as the system architecture, the presence of a web interface, the language used, supported platforms, the need to work on a managed node. The technical requirements necessary for the correct functioning of the configuration management system and the corporate mail server were determined in the work. The created corporate mail system is designed for the highest possible level of security and confidentiality of business correspondence, convenient and fast interaction of employees, as well as monitoring and managing the university's email traffic, protection from spam and external attacks. This method of organizing corporate mail using a configuration management system simplifies the work of configuring and deploying a mail server, optimizing the work of technical personnel and the educational institution as a whole.

Keywords: mail server, corporate mail, configuration management tool, iReaMail, Ansible

Graphical annotation (Графическая аннотация)



Введение. Деятельность любой современной организации сложно представить без использования универсального сервиса по обмену данными между пользователями корпоративной сети – электронной почты. Основное предназначение электронной почты в любой организации – связь пользователей друг с другом с целью передачи документации, распоряжений и прочего документооборота. Ежедневно отправляются сотни сообщений, без которых невозможно осуществить продуктивную деятельность компании. В связи с этим наличие защищенного почтового сервера на вычислительных мощностях образовательной организации является крайне важным, поскольку это предотвратит потерю важной документации и распространение вредоносного программного обеспечения.

Развертывание корпоративной почтовой системы можно и нужно автоматизировать с целью экономии времени персонала технических отделов организаций. Одной из самых наиболее стабильных систем управления конфигурациями и автоматизации программного обеспечения и операционных систем является Ansible [3].

Значительные затраты на реализацию проекта на собственном почтовом сервере в итоге гарантируют ряд преимуществ, к которым можно отнести, прежде всего, оперативный мониторинг и управление работой почтовым сервером, собственную систему хранения, возможность создания персональных корпоративных почтовых ящиков сотрудников.

Крайне важно выбрать такое средство управления конфигурациями, которое будет отвечать всем требованиям конкретного сервера с точки зрения настроек всевозможных опций и компонентов. На рынке программного обеспечения каждый программный комплекс предлагает свой функционал, направленный на обеспечение автоматизации процессов и быстрое развертывание узлов и приложений.

Все вышесказанное подтверждает актуальность создания корпоративной почтовой системы с использованием системы управления конфигурациями, что и является целью данной статьи.

Оценка возможности развертывания почтового сервера. На данный момент существует четыре основных способа организации корпоративной почты.

Самый простой вариант – создать каждому сотруднику собственный бесплатный аккаунт на одной из популярных почтовых служб (обычно такой почтовый ящик уже ранее создан). Это не требует каких-либо финансовых вложений от компании, что является несомненным плюсом. Однако такой выбор имеет несколько ограничений, влияющих на продуктивность коммуникации: низкая производительность и перебои работы за счет высокой нагрузки сервера, отсутствие оперативной техподдержки, постоянная демонстрация таргетированной рекламы, риски утечки данных, ограничение на размер отправляемой корреспонденции (не более 2 Гб), вероятность утраты доступа к почте (взлом, блокировка непосредственно платформы), утрата положительного имиджа компании и др.

Использование собственного доменного имени на бесплатных (условно-бесплатных) сервисах – альтернативный вариант, пользующийся популярностью у компаний с небольшим числом сотрудников. Как правило, руководитель компании имеет зарегистрированный почтовый аккаунт на таком домене (например, company.bizml.ru в экосистеме Mail.ru). B2B-решение реализовано как для настольной версии, так и мобильной и включает почтовый сервис, список контактных адресов, расписание событий. Зарегистрированный сотрудник может использовать свою учетную запись для непосредственного онлайн-общения в мессенджере, а также размещать различные документы в облачное хранилище. В этом варианте к возможностям первого способа создания почтового сервиса добавляются платные услуги (например, расширение объема дискового пространства корпоративного почтового ящика).

Чаще всего компании среднего и малого бизнеса предпочитают организовывать корпоративную почту, используя предоставляемые услуги хостинг-провайдеров: использование непосредственно цифрового пространства хостинга, VDS или выделенный сервер, который может арендовать организация. Третий способ предоставляет более широкие возможности по сравнению с двумя предыдущими вариантами. Например, использование не только веб-интерфейса для доступа к почте (браузерная версия), но и локальных почтовых клиентов (настольная версия). Сложности работы в таком режиме чаще всего связаны с аппаратной частью: невозможность самостоятельно устранить физические неполадки оборудования из-за ограничений доступа к серверной части провайдера, запрет на персонализированные настройки сервера. Другой проблемой может стать уменьшение выделенной памяти за счет расширения штата сотрудников и увеличения объемов корреспонденции и, как следствие, необходимость выбора нового тарифного плана более высокой стоимости.

Наиболее беспроигрышным можно считать четвертый подход – организация почтового сервера на базе собственных вычислительных мощностей.

Для этого необходимо наличие специального оборудования и определенных технических возможностей:

- скоростной и высоконадёжный доступ к сети Интернет (для повышения стабильности системы желательно иметь не менее двух провайдеров);
- физический сервер под управлением операционной системы, как правило, Windows Server или Linux;
- комплекс программ для обслуживания почтовой системы.

Компания, исходя из своих ресурсов и возможностей (род деятельности, объемы производства, финансово-хозяйственная деятельность, численность работников и пр.), может прибегнуть к различным вариантам организации корпоративной почты (рис. 1).

Анализ рассмотренных выше вариантов создания корпоративной почтовой системы позволяет сделать вывод о выборе четвертого способа как оптимального для образовательного учреждения. Данный выбор позволит обеспечить полноценную систему мониторинга работоспособности почтового сервера, организовать управление дисковым пространством и резервным копированием данных.

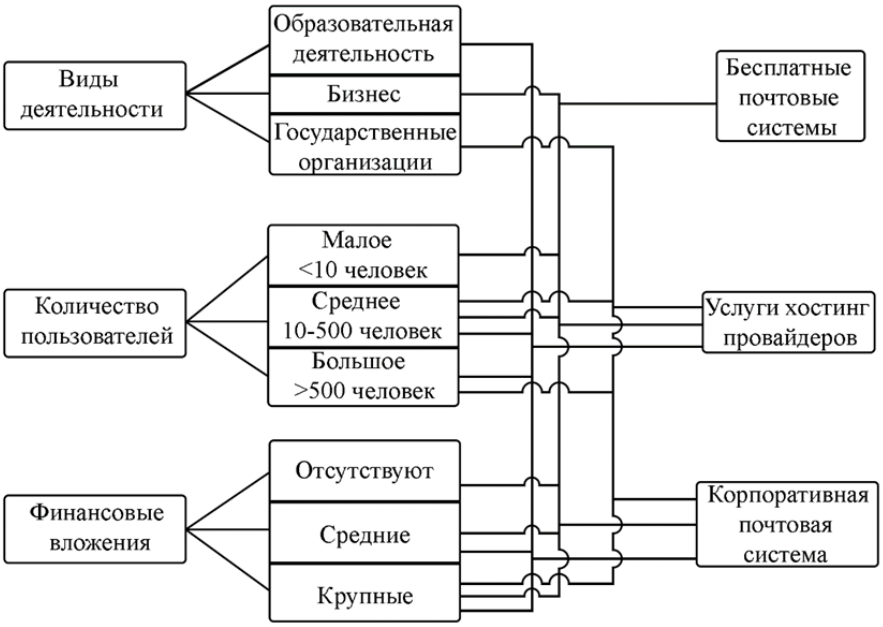


Рисунок 1 – Выбор способа организации корпоративной почты

iRedMail – бесплатное решение с открытым кодом (Open source) для создания почтовых серверов, которое представляет собой совокупность сценариев для автоматизации инсталляции и настройки структурных элементов, обеспечивающих полноценную работу почтового сервера организации [6]. При развертывании почтовых серверов iRedMail по сравнению с ручной настройкой компонентов дает возможность уменьшить временные затраты и не требует учета специфических тонкостей Linux-дистрибутивов.

В состав входят такие инструменты, как веб-интерфейс для администрирования iRedMailamdin, веб-интерфейс для доступа к электронной почте, спам-фильтры, антивирус, а также IMAP-сервер для доступа к почтовым ящикам пользователей.

iRedMail имеет общую адресную книгу, что позволяет отправлять сообщения внутри организации по названию структурного подразделения или по сведениям о сотруднике (фамилия, имя, инициалы, логин…). Встроенная функция формирования списков исключений позволяет запретить отправку писем некоторым пользователям без непосредственного взаимодействия с аккаунтом пользователя (так называемые белые и черные списки).

В почтовый сервер включен антивирус Clamav, который способствует предотвращению получения вредоносного ПО пользователем корпоративной почты, тем самым обеспечивая безопасность рабочих станций организации. Помимо этого, в состав почтовой системы входит модуль SpamAssassin, который обеспечивает защиту от рекламных и фишинговых сообщений.

Почтовый сервер iRedMail способен переадресовывать почтовые сообщения. Можно настроить сценарий таким образом, чтобы сообщение было отправлено не только на указанный адрес, но и на дополнительный как копия. То есть письмо получат оба пользователя, при этом ни один из них не будет знать о дублировании сообщения кому-либо еще (например, при рассылке информационных оповещений партнерам или потенциальным клиентам).

Также iRedMail поддерживает Throttle-лимиты для установки ограничений на количество отправленных писем, их максимальный объем в указанный период времени (минута, час, сутки) с конкретных почтовых адресов или со всех существующих адресов почтового сервера.

Системы управления конфигурациями. Централизованное управление различными имеющимися операционными системами и пакетами прикладных программ, установленных на оборудовании организации, обеспечивают системы управления конфигурациями (CM-системы). Такие системы автоматизируют рутинные цепочки действий администратора по настройке конфигурации серверов, облегчая тем самым обслуживание постоянно растущих мощностей IT-инфраструктуры предприятия.

На сегодняшний момент на рынке программного обеспечения существует множество различных систем управления конфигурациями, имеющими однотипный функционал, при этом обладающими принципиальными различиями [2].

Был проведен критериальный анализ наиболее популярных СМ-систем по следующим параметрам (табл. 1):

- архитектура СМ-системы;
- язык скриптов;
- платформы для управления;
- необходимость работы непосредственно на управляемом узле.

Таблица 1 – Характеристики СМ-систем

СМ-системы / Характеристики	Ansible	Chef	Puppet	SaltStack
Архитектура СМ-системы				
централизованная	+	+	+	+
слабо распределенная		+	+	+
Язык описания сценариев				
YAML	+			+
Python	+			+
на основе Ruby		+	+	
Платформы для управления				
Unix, MacOC, Windows	+++	+++	+++	+++
Работа с клиентом				
клиент отсутствует	+			
ручная настройка			+	+
автоматическая установка		+		

Многие инструменты кода инфраструктуры, включая рассмотренные, используют декларативные языки. Код определяет желаемое состояние инфраструктуры, например, какие пакеты и учетные записи пользователей должны находиться на сервере или сколько ресурсов памяти (ОЗУ/ЦП) он должен иметь [9].

Знание языка программирования влияет на выбор и дальнейшее использование СМ-системы. Для работы с системой Puppet в штате сотрудников должен быть как минимум middle-разработчик со знанием тонкостей программирования на языке Ruby [8].

Начинающему администратору будет достаточно сложно освоить работу со сценариями (скриптами) системы Chef, а также от него потребуются знания особенностей сетевых настроек (ошибки в настройках могут привести не только к потере соединения, но и вплоть до потери данных и выхода из строя сетевого оборудования) [4].

Описанные выше СМ-системы рассчитаны в большей мере на наличие в штатном расписании программиста, а системы Ansible и SaltStack рассчитаны непосредственно на системных администраторов, знание языков программирования у которых приветствуется, но необязательно.

Управление облачными инфраструктурами с гибкими конфигурациями администраторы предпочитают выполнять в хорошо масштабируемой среде SaltStack [10]. Однако веб-интерфейс менее функционален по сравнению с другими вариантами, а инструментарий для детализации отчетов по работе сервера и узлов ограничен.

Большинство инструментов декларативной инфраструктуры расширяют свои языки, добавляя возможности императивного программирования. Например, Ansible добавляет в YAML циклы и условные выражения. Многие инструменты инфраструктуры не только декларативны, но и используют собственный предметно-ориентированный язык (DSL), разработанный для моделирования конкретной области (в частности инфраструктуры). Это упрощает написание и понимание кода. Например, у Ansible, Chef и Puppet есть DSL для настройки серверов. Их языки предоставляют конструкции для таких понятий, как пакеты, файлы, службы и учетные записи пользователей.

СМ-система Ansible подходит как начинающему, так и опытному системному администратору [3]. Отличительной особенностью является не только архитектура системы, не требующая локальных агентов на управляемых узлах, но и возможность написания модулей практически на любом языке (как правило, используется язык Python). Управление конфигурацией узла с предустановленным Python позволяет не использовать дополнительное программное обеспечение.

Система Ansible может быть использована в одном из двух вариантов: Open Source и Enterprise. Корпоративная версия Ansible Tower имеет дополнительную графическую оболочку, реализующую автозапуск ряда сценариев и некоторых дополнительных прикладных задач.

Развертывание корпоративной системы. Масштабы движения внешних и внутренних информационных потоков (как получаемых, так и передаваемых) растут с каждым днем. Доступность

и своевременное осуществление доставки актуальных данных в структурные подразделения организации имеют немаловажное значение для оперативного принятия управленческих решений.

Немаловажную роль среди всевозможных способов коммуникации в работе компании играет электронная почта как часть электронного обмена информацией. Расширение инфраструктуры организации влечет за собой увеличение объемов электронных почтовых сообщений, отправляемых сотрудниками (например, информирование о предстоящих событиях), а также объемов входящих писем (например, регулярные рассылки).

Корпоративная почтовая система на собственном домене дает возможность организовать поддержку не только внутренней переписки (взаимообмен файлами), но и с внешними источниками информации (сообщения от деловых партнёров или потенциальных клиентов). Этот вариант позволяет решить проблему учета входящих сообщений, которые могут быть нежелательными и являться навязчивой электронной рассылкой, спамом, фишинговой атакой.

В рамках научно-исследовательской темы «Цифровая среда» на основе договора с НКО «Фонд содействия развитию науки и образования» развернута система управления конфигурациями Ansible с целью совершенствования функционирования и регулирования работы корпоративной сети ФГБОУ ВО «Брянский государственный университет имени академика И.Г. Петровского (БГУ)» [1].

На рисунке 2 представлен перечень задач, решаемых специалистом по настройке и поддержке ИТ-инфраструктуры компании (системным администраторам), а также реализуемых СМ-системой в целом.

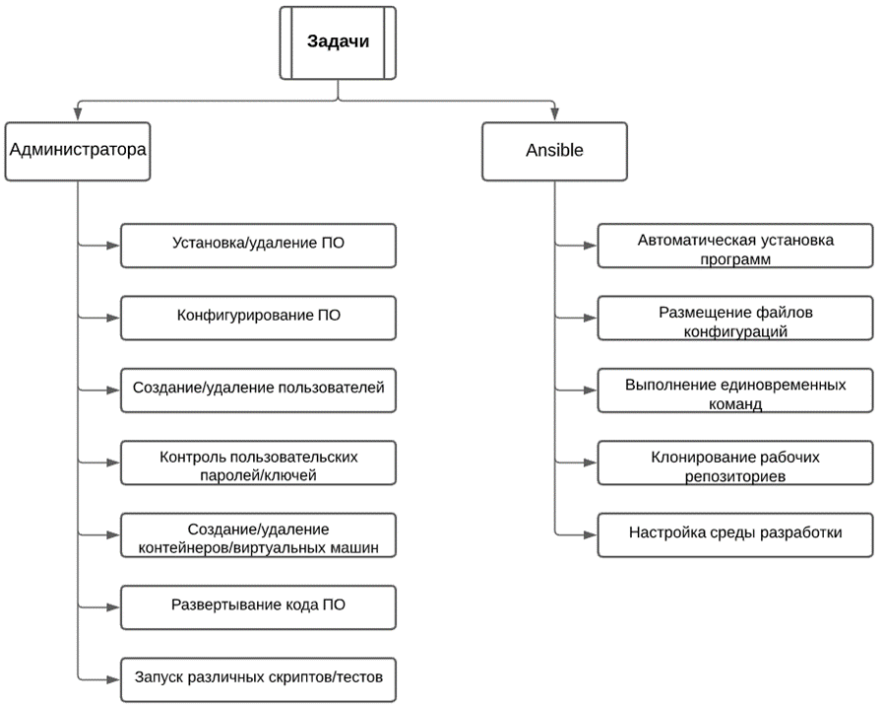


Рисунок 2 – Задачи по настройке и поддержке ИТ-инфраструктуры компании

При построении локальной вычислительной сети вуза использовался макет топологии «звезда», основой которого послужит центр хранения и обработки данных (ЦХОД). В его состав вошли многопроцессорные сервера (7 серверов с 10 процессорами и 54 ядрами, объемы памяти: ОП – 248 Гбайт, ДП – 26 Тбайт). Вычислительная мощность составляет 300 Гфлопс.

Развертывание системы Ansible выполнялось поэтапно: сначала установка и настройка системы, а затем создание сценариев работы. Для управления конфигурациями сетевого оборудования (приобретены маршрутизаторы и сетевые коммутаторы MikroTik) в системе Ansible изначально требуется задать ряд параметров:

1. Добавить группу [routers] со списком подключенных маршрутизаторов в файл инвентаризации. Это позволит упростить и унифицировать выполнение задач за счет обращения к группе устройств без необходимости перечисления адресов каждого устройства в отдельности.

2. Настроить при необходимости пользовательские параметры подключения к устройствам. Это позволит выполнять задачу, привлекая как одно конкретное устройство, так и группу или всё оборудование.

Каждый Ansible-скрипт (модуль) реализует выполнение задачи с заданными параметрами в одном из двух режимов: ad-hoc команда или playbook (плейбук). В первом случае выполняются разовые действия с узлами или группами нодов, как правило, через консоль [7].

Решение нестандартных задач (одной или сразу нескольких) осуществляется через плейбуки – наборы команд по проверке выполняемости определенных условий (рис. 3). Playbook позволяет определить группу хоста, в которой выполняются задачи, и может содержать различную информацию об управляемом узле [11].

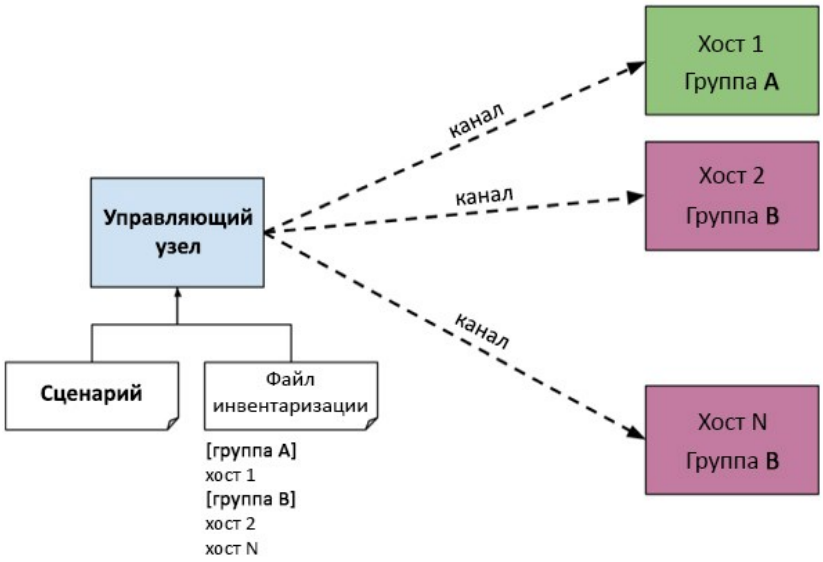


Рисунок 3 – Выполнение Ansible-скрипт

Мониторинг сетевого оборудования включает в себя следующие варианты:

1. Получение сведений о конкретном устройстве выполняется с помощью плейбука data_get.yml, результат работы которого отобразится в терминале командной строки (листинг 1).

Листинг 1

```
- - -
- name: Getting data from routers
  hosts: routers
  gather_facts: false
  tasks:
    - name: system command
      routeros_command:
        commands: /system routerboard print
      register: system_print
    - name: debug print
      debug: var=system_print.stdout_lines
```

2. Отправка файла на устройство выполняется с помощью плейбука file_set.yml, например, загрузить скрипт для автоматического обновления роутера (листинг 2).

Листинг 2

```
---
- name: Set file to router
  hosts: 10.0.3.203
  gather_facts: false
  tasks:
    - name: file upload
      net_put:
        src: ./spw_script
        protocol: sftp
    - name: File demonstate after
      routeros_command:
        commands: /file print
      register: file_after
    - name: debug print after
```

3. Получение файла с устройства выполняется с помощью плейбука file_get.yml, например, создание резервной копии на устройстве и ее отправка в указанную директорию (листинг 3).

Листинг 3

```
- - -
- name: Save file from router
  hosts: 172.16.16.12
  gather_facts: false
  tasks:
    - name: backup-file create
      routeros_command:
        commands: /system backup save name=backup_spw
    - name: backup-file download
      net_get:
        src: backup_spw.backup
            protocol: sftp
```

Подготовленными плейбуками (Ansible-сценариями) могут воспользоваться администраторы корпоративных сетей на базе системы управления конфигурациями Ansible для сокращения рутинных работ по обслуживанию сетевого оборудования.

Код конфигурации сервера, как и любой другой код, со временем превращается в беспорядок. Системы управления конфигурациями позволяют не только избежать этого, но и структурировать код по разным блокам (модулям). Администратор может организовывать, редактировать и выпускать эти модули по отдельности или группами.

Ansible является мощным инструментом, обеспечивающим надежное управление разветвленными сетями с нодами под управлением различных операционных систем.

Развертывание почтового сервера iRedMail с использованием системы управления конфигурациями требует учета ряда технических требований:

- максимально полный охват IT-инфраструктуры организации (сервера, сетевое оборудование, стабильное высокоскоростное интернет-соединение);
- соотнесение почтового сервера с доменным именем организации;
- инсталляция функциональных пакетов для бесперебойной работы почтового сервера;
- установка операционной системы Linux Ubuntu (версия 18 и выше);
- наличие установленного Python не ниже версии 2.4 (входит в Linux-дистрибутив).

Установленные пакеты реализуют следующий функционал:

- работа с почтовым ящиком в браузере (модуль roundcube);
- отправка электронной почты (исходящая почта) по протоколу smtp (модуль postfix);
- получение сообщений (входящая почта) по протоколу pop3/imap (модуль dovecot);
- сбор сведений о текущем состоянии (модуль awstats);
- фильтрация вирусов и спама (модули amavisd, spamassassin, clamav, greylist).

Осуществление системного администрирования возможно из любой сети с помощью вебпанели управления почтовым сервером iRedAdmin (страница /iredadmin). Администратор может постоянно мониторить работу серверов и управляемых узлов, выполнять своевременное обновление необходимого программного обеспечения и оперативно решать возникающие проблемы. Используя панель управления, можно создавать новых пользователей с привязкой почтовых ящиков и редактировать параметры существующих почтовых ящиков.

Зарегистрированный пользователь может войти в созданный почтовый ящик для отправки и получения сообщений через браузер со страницы mx1.brgu.ru или воспользоваться настольным email-приложением (например, The Bat!, Microsoft Outlook, TouchMail).

Внедрение корпоративной почтовой системы позволило реализовать и обеспечить устойчивое функционирование, в том числе:

- организацию менеджмента ИТ-инфраструктуры университета;
- своевременный мониторинг всего парка сетевого оборудования;
- минимизацию временных затрат на установку и настройку конфигураций системы и управляемых нодов;
- автоматизацию рутинных работ за счет приведения их к единой форме и централизованного исполнения;
- снижение влияния субъективных факторов;
- применение средств защиты электронной почты для повышения ее безопасности и конфиденциальности;
- настройку уровней доступа и привилегий для различных категорий пользователей;
- организацию оперативного обмена сообщениями;
- мониторинг и управление трафиком.

Заключение. Управление сетевой инфраструктурой крупного образовательного учреждения требует разработки множества сценариев установки и настройки (определение физических и виртуальных управляемых нодов сети, инсталляция необходимого программного обеспечения, настройка сервисов и служб и пр.). Использование динамичного характера современной инфраструктурной платформы Ansible, уделение внимания тестированию и согласованности позволили существенно уменьшить временные затраты и увеличить результативность действий по конфигурированию и мониторингу серверного и сетевого оборудования корпоративной сети ФГБОУ ВО «Брянский государственный университет имени академика И.Г. Петровского».

Корпоративный адрес – это визитная карточка организации. Все отправляемые почтовые сообщения имеют адрес в формате user@brgu.ru. Это, с одной стороны, ассоциирует его с образовательной организацией и повышает ее узнаваемость. Кроме того, позволяет увеличить процент доставленных и прочитанных писем от компании (по сравнению с отправляемыми с общих доменов письмами, которые, вероятнее всего, попадают в спам).

Создание корпоративной почтовой системы обеспечило полный контроль над работой почтового сервера, легкую масштабируемость объема дискового хранилища, неограниченное количество пользователей корпоративной почты, мониторинг внутренней активности сотрудников (например, время отклика на полученное сообщение).

Библиографический список

1. Беднаж, В. А. Реализация современной концептуальной модели исследования задач в области прикладных Интернет-технологий / В. А. Беднаж, Н. А. Иванова, С. Е. Саланкова // Современные наукоемкие технологии. – 2018. – № 10. – С. 22–26.

2. Савватеев, М. Е. Возможности систем управления конфигурациями при развертывании и настройке почтовой системы в рамках деятельности образовательной организации / М. Е. Савватеев, А. М. Селезнева // Университет на пути к новому качеству науки и образования : Национальная научно-практическая конференция с международным участием. – 2020. – 5 с.

3. Ansible Documentation. – Режим доступа: http://docs.ansible.com, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 02.03.2021).

4. Chef Progress. – Режим доступа: https://www.chef.io., свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 18.03.2021).

5. Hochstein, Lorin. Ansible: Up and Running. Automating Configuration Management and Deployment the Easy Way / Hochstein Lorin, Moser René. – Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2017. – 429 p.

6. iRedMail – Open Source Mail Server Solution. – Режим доступа: https://www.iredmail.org, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 21.03.2021).

7. Keating, Jesse. Mastering Ansible – Third Edition / Keating Jesse. – 3rd ed. – Packt Publishing Ltd., 2019. – 412 p.

8. Make infrastructure actionable, scalable and intelligent. – Режим доступа: https://puppet.com, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 18.04.2021).

9. Morris, Kief. Infrastructure as Code / Morris Kief. – 2nd ed. – Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2021. – 430 p.

10. Salt Project. – Режим доступа: https://saltproject.io, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 18.03.2021).

11. Unleashing full potential of Ansible Framework: university labs administration / P. Masek, M. Stusek, Ja. Krejci, K. Zeman, J. Pokorny, M. Kudlacek // Conference of open innovations association (FRUCT). – 2018. – № 22. – P. 144–150.

References

1. Bednazh, V. A. Realizatsiya sovremennoy kontseptualnoy modeli issledovaniya zadach v oblasti prikladnykh Internet-tekhnologiy [Implementation of a modern conceptual model for the study of problems in the field of applied Internet technologies]. *Sovremennye naukoemkie tekhnologii* [Modern High Technologies], 2018, no. 10, pp. 22–26.
2. Savvateev, M. E. Vozmozhnosti sistem upravleniya konfiguratsiyami pri razvertyvanii i nastroyke pochtovoy sistemy v ramkakh deyatelnosti obrazovatelnoy organizatsii [The Capabilities of configuration management when you deploy and configure the mail system in the framework of the activities of an educational institution]. *Universitet na puti k novomu kachestvu nauki i obrazovaniya : Natsionalnaya nauchno-prakticheskaya konferentsiya s mezhdunarodnym uchastiem* [The University on the way to a new quality of science and education"National scientific-practical conference with international participation], 2020. 5 p.
3. *Ansible Documentation.* Available at: http://docs.ansible.com (accessed 02.03.2021).
4. *Chef Progress.* Available at: https://www.chef.io (accessed 18.03.2021).
5. Hochstein, Lorin, Moser, René. *Ansible: Up and Running. Automating Configuration Management and Deployment the Easy Way.* Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2017. 429 p.
6. *iRedMail – Open Source Mail Server Solution.* Available at: https://www.iredmail.org (accessed 21.04.2021).
7. Keating, Jesse. *Mastering Ansible – Third Edition*. 3rd ed. Packt Publishing Ltd., 2019. 412 p.
8. *Make infrastructure actionable, scalable and intelligent.* Available at: https://puppet.com (accessed 18.04.2021).
9. Morris, Kief. *Infrastructure as Code.* 2nd ed. Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2021. 430 p.
10. *Salt Project.* Available at: https://saltproject.io (accessed 18.03.2021).
11. Masek, P., Stusek, M., Krejci, Ja., Zeman, K., Pokorny, J., Kudlacek, M. Unleashing full potential of Ansible Framework: university labs administration. *Conference of open innovations association (FRUCT)*, 2018, no. 22, pp. 144–150.

DOI 10.54398/2074-1707_2022_1_26
УДК 504

СИСТЕМНЫЙ АНАЛИЗ ОСНОВНЫХ ЭКОЛОГИЧЕСКИХ УГРОЗ ДЛЯ ОКРУЖАЮЩЕЙ СРЕДЫ ПРИ ЭКСПЛУАТАЦИИ ГАЗОВЫХ СКВАЖИН

Статья поступила в редакцию 15.07.2021, в окончательном варианте – 13.02.2022.

Амшинов Ника Мерабович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

аспирант, ORCID: 0000-0002-7107-1610, e-mail: amshinov.nika@yandex.ru

Ажмухамедов Искандар Маратович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

доктор технических наук, декан факультета цифровых технологий и кибербезопасности, профессор кафедры информационной безопасности, ORCID: 0000-0001-9058-123X, e-mail: aim_agtu@mail.ru

Статья посвящена анализу основных экологических угроз, оказывающих негативное воздействие на окружающую среду, при эксплуатации газовых скважин. Рассмотрены возможные осложнения, возникающие на газовых скважинах на этапе эксплуатации, а также основные негативные экологические воздействия на окружающую среду. Проведен анализ работ, посвященных возможным осложнениям на скважинах, приводящих к негативному воздействию на окружающую среду, и способам их устранения. Систематизированы основные осложнения, приведенные в источниках. Отмечено, что последствия, вызванные осложнениями либо мероприятиями по ликвидации осложнений, в свою очередь, приводят к негативному экологическому воздействию на окружающую среду, что впоследствии может грозить штрафами для компании и затратами на ликвидацию последствий от негативного воздействия на экологию. Выделены группы экологических происшествий по среде, на которую оказывается воздействие. Построена онтологическая схема загрязнения окружающей среды при эксплуатации скважин. Введен критерий эффективности, а именно «минимальный риск для компании», позволяющий минимизировать возможные экологические и экономические риски от принимаемых решений по ликвидации осложнений. Данный критерий представляет собой сумму рисков, связанных с технологическими осложнениями на газодобывающих скважинах и экологическими происшествиями, вызванными данными осложнениями.

Ключевые слова: системный анализ, экологические угрозы, газовые скважины

SYSTEM ANALYSIS OF THE MAIN ENVIRONMENTAL THREATS TO THE ENVIRONMENT DURING THE OPERATION OF GAS WELLS

The article was received by the editorial board on 15.07.2021, in the final version – 13.02.2022.

***Amshinov Nika M.*,** Astrakhan State University, 20a Tatischev St. Astrakhan, 414056, Russian Federation,

postgraduate student, ORCID: https://orcid.org/0000-0002-7107-1610, e-mail: amshinov.nika @yandex.ru

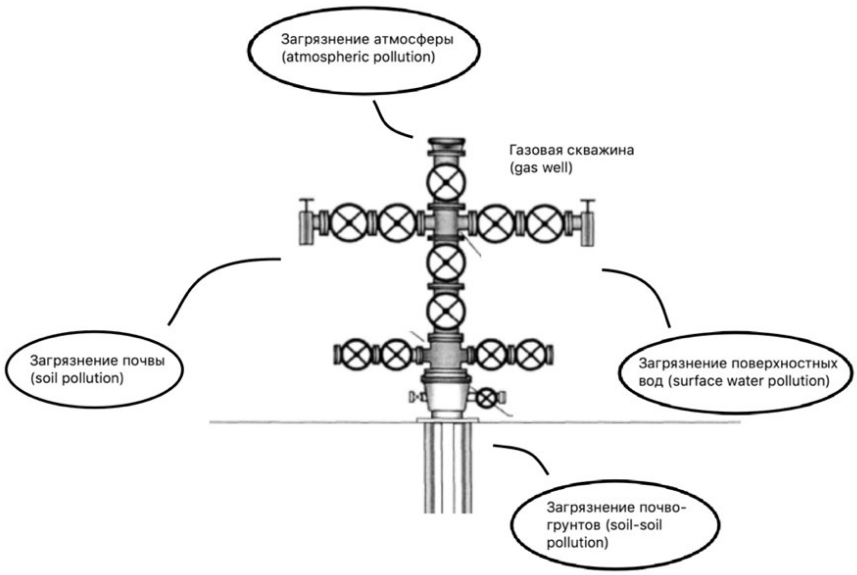
***Azhmukhamedov Iskandar M.*,** Astrakhan State University, 20a Tatischev St. Astrakhan, 414056, Russian Federation,

Doct. Sci. (Engineering), Dean of the Faculty of Digital Technologies and Cybersecurity, Professor of the Department of Information Security and Digital Technologies, ORCID: https://orcid.org/0000-0001-9058-123X, e-mail: aim_agtu@mail.ru

The article is devoted to the analysis of the main environmental threats that have a negative impact on the environment during the operation of gas wells. Possible complications arising at gas storage facilities at the operational stage, as well as the main negative environmental impacts on the environment are considered. The analysis of works devoted to possible complications at gas wells leading to a negative impact on the environment and ways to eliminate them has been carried out. The main complications listed in the sources are systematized. It is noted that the consequences caused by complications or measures to eliminate complications, in turn, lead to a negative environmental impact on the environment, which can subsequently threaten fines for the company and the costs of eliminating the consequences of a negative impact on the environment. Groups of environmental incidents are identified by the environment that is affected. The ontological scheme of environmental pollution during the operation of wells is constructed. An efficiency criterion has been introduced, namely, "minimum risk for the company", which allows minimizing possible environmental and economic risks from decisions taken to eliminate complications. This criterion is the sum of the risks associated with technological complications at gas producing wells and environmental accidents caused by these complications.

Keywords: system analysis, environmental threats, gas wells

Graphical annotation (Графическая аннотация)



Введение. Обеспечение экологической безопасности является важной проблемой при эксплуатации опасных производственных объектов (ОПО), связанных с добычей, транспортировкой и переработкой углеводородов (УВ). Негативное экологическое воздействие на окружающую среду (ОС) стало стремительными темпами возрастать в связи с увеличением доли используемых пожаро-, взрыво-, химически опасных технологий, являющихся основными источниками загрязнения природы. Несмотря на то, что разработано множество нормативных документов федерального и отраслевого уровня, направленных на обеспечение экологической безопасности, воздействие на природу со стороны ОПО остается на очень высоком уровне.

На сегодняшний день ведется активная работа по созданию и внедрению систем, предназначенных для мониторинга состояния окружающей среды и снижения риска возникновения аварийных ситуаций на ОПО. Особенно остро стал вопрос несвоевременного предупреждения экологических происшествий в связи с последними крупными авариями в Норильске (29 мая 2020 г. – крупнейшая утечка нефтепродуктов из резервуара [22]) и в Оренбургской области (23 февраля 2021 г. – взрыв федерального магистрального газопровода [23]). К сожалению, в случае экологических катастроф часто невозможно точно определить, какой вред нанесла утечка углеводородов окружающей среде и возможно ли полное восстановление экосистемы.

Газодобывающие предприятия также относятся к ОПО, а одним из самых опасных объектов данного производства является газодобывающая скважина (ГДС). Так, например, на Астраханском газоконденсатном месторождении (АГКМ) идет добыча газа с высоким содержанием кислых компонентов (углекислого газа, сероводорода, диоксида серы), что приводит к негативному воздействию на окружающую среду Прикаспийского региона.

Обеспечение охраны ОС становится все более сложной задачей в связи с ростом добычи УВ, а также проведением геологоразведки на территориях с ограниченным режимом природопользования.

Исходя из вышеизложенного, совершенствование существующих и разработка новых систем, которые будут предупреждать экологические происшествия и обеспечивать экологическую безопасность окружающей среды при эксплуатации газовых скважин, является весьма актуальным. Под термином «экологические происшествия» понимается негативное экологическое воздействие ГДС на окружающую среду в связи с осложнениями.

Жизненный цикл ГДС состоит из следующих этапов: проектирование, строительство, эксплуатация, ликвидация. Наиболее серьезные негативные экологические воздействия на окружающую среду происходят на этапах, изображенных на рисунке 1. На каждом из приведенных этапов могут возникать осложнения. При этом под термином «осложнение» понимаются затруднения при эксплуатации ГДС, которые могут привести к негативным экологическим последствиям для окружающей среды [21].

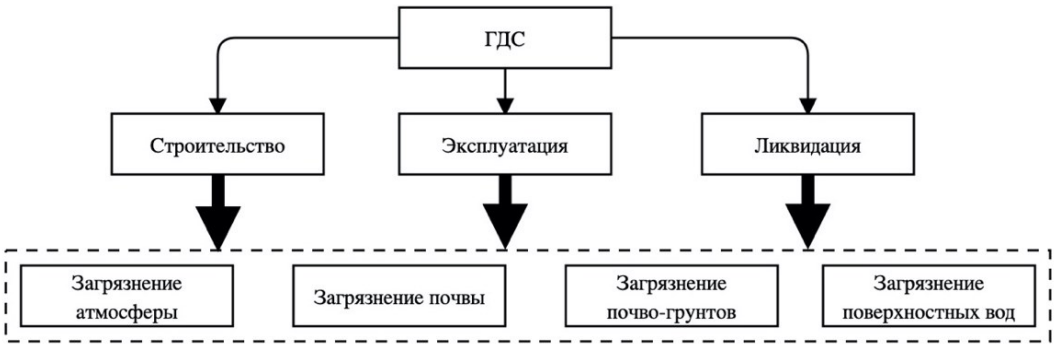


Рисунок 1 – Негативные экологические воздействия на этапах жизненного цикла ГДС

Проблемами возникновения осложнений и их негативного воздействия на экологию на этапе бурения ГДС занимались Ю.Г. Безродный, И.Ю. Макаренкова, А.И. Булатов и др. [11–13]. Проблеме негативного экологического воздействия (НЭВ) на этапе ликвидации ГДС посвящены работы И.И. Агадулина, О.Я. Зорина, В.Г. Султанова и др. [16–18]. Имеются также работы, посвященные проблемам, возникающим в процессе эксплуатации ГДС (Ю.П. Коротаев, Ш.К. Гиматудинов, З.С. Алиев и др. [1–2, 5]). Однако данная проблема на сегодняшний день остается недостаточно проработанной.

Постановка задачи. В этой связи необходимо разработать методику оценки рисков возникновения осложнений, влекущих негативное экологическое воздействие на окружающую среду в процессе эксплуатации газовых скважин. Это, в свою очередь, требует системного анализа влияния технологических параметров, характеризующих процесс эксплуатации ГДС, на вероятность возникновения различных осложнений, которые приводят к загрязнению атмосферы, почв, подземных и поверхностных вод и т. п.

Системный анализ основных экологических угроз для окружающей среды при эксплуатации газовых скважин. Рассмотрим подробнее возможные экологические риски, которые могут возникнуть в процессе эксплуатации газовых скважин.

НЭВ на окружающую среду со стороны газовых скважин возникают в результате плановых (прогнозируемых) и внеплановых (непрогнозируемых) осложнений, которые могут привести к экологическим происшествиям.

Изучению возможных осложнений на газовых скважинах, приводящих к негативному воздействию на ОС, и способам их устранения посвящены работы [1–9].

В работе [1] рассмотрено влияние жидкости в стволе и на забое на работу газовых и газоконденсатных скважин, а также приведены основные результаты исследований по данному вопросу. Под термином «ствол» скважины понимается пространство в массиве горных пород, ограниченное контурами скважины, т. е. ее устьем, стенками и забоем, а «забой» – нижняя часть скважины, вскрывающая продуктивный пласт. В процессе работы газовых и газоконденсатных скважин происходит конденсация водяных паров и тяжелых углеводородов из-за дросселирования и теплообмена газа со стенками скважины. Содержание конденсационной воды и конденсата может быть незначительным, но в процессе эксплуатации спустя время может скопиться значительное количество жидкости (от 0,1 до 40 кг на 1000 м^3 газа), а высота столба жидкости будет зависеть от дебита газа, давления и конструкции ствола скважины.

Практика эксплуатации газовых и газоконденсатных месторождений подтверждает факт, что в большинстве газовых и газоконденсатных месторождений в призабойной зоне пласта (ПЗП) и в стволе скважин конденсируются водяные пары и жидкие УВ. Помимо конденсационной воды, в пласте в зависимости от его строения (конфигурации и размеров пор) может содержаться от 5 до 50 % остаточной воды (считая от свободной пористости), которая также может выноситься вместе с газом.

Наряду с этим, эксплуатация газовых скважин осложняется из-за образования конуса воды вследствие притока подошвенных вод в связи с большими дебитами скважин. Вода поступает в ствол скважины, что снижает дебиты скважин по газу и увеличивает нагрузку на наземные сооружения. В этой связи авторами были выделены две группы скважин в зависимости от количества жидкости, дебита газа, давления, конструкции ствола и забоя. К первой группе отнесены скважины, в которых не происходит вынос жидкости на поверхность, а ко второй – скважины, в которых происходит вынос жидкости на поверхность. Для устранения рассмотренных осложнений на ГДС требуется проведение ремонтных работ, которые могут привести к загрязнениям подземных вод и почво-грунтов.

В [2] рассмотрены химические, механические и тепловые методы воздействия на ПЗП, основанные на искусственном увеличении проводимости пород, в результате которых возможно повышение производительности скважин. Каждому из методов дана оценка эффективности их применения.

Одним из распространенных методов является кислотная обработка пласта. Суть метода заключается в закачке в пласт определенных кислот (соляной, серной или фтороводородной) с целью образования каналов разъедания, которые соединяют забой скважин с насыщенными газом участками пласта. Наряду с хорошей результативностью данного метода существует ряд основных недостатков, а именно:

1. Кислота вызывает коррозию внутрискважинного оборудования (ВСО), поэтому с целью ее предотвращения добавляют ингибиторы коррозии.

2. Большое количество вредных примесей приводит к образования новых нерастворимых соединений, вызывающих кольматацию ПЗП (для минимизации риска образования таких соединений добавляют такие стабилизаторы, как хлористый барий, уксусная кислота и др.).

Следующий метод – гидравлический разрыв пласта (ГРП). В результате данного метода происходит образование, под воздействием давления, трещин в связи с нагнетаемой в скважину плохо фильтрующейся жидкости. Основными этапами процесса можно отметить:

1) закачка жидкости в пласт для образования трещин, которые заполняются крупнозернистым песком;

2) нагнетание жидкости-песконосителя;

3) закачка жидкости для продавливания песка в скважину. Данный метод является эффективным, но требует серьезных трудозатрат и значительных средств, а также материалов для реализации.

Еще одним методом является гидропескоструйная перфорация. Суть метода в перфорации гидропескоструйным перфоратором, спускаемым в скважину на насосно-компрессорных трубах (НКТ), обсадной колонны и породы, что позволит увеличить приток газа. К сожалению, данный метод не всегда может быть эффективен в связи с необратимыми механическими изменениями в строении пород из-за снижения пластового давления в процессе добычи и ухудшения фильтрационных свойств под влиянием возрастающего эффективного вертикального напряжения.

Четвертым методом является теплофизическое воздействие на ПЗП. Данным методом происходит удаление парафина, смол и солей. Также периодически прогреваются породы пласта вокруг скважины, благодаря чему достигается сохранение фильтрационных свойств пласта.

Пятый метод – импульсно-ударное и вибрационное воздействие. С помощью сильных ударных волн, генерирующихся при взрыве глубинных зарядов (или бомб), повышается проводимость пласта из-за образования сети трещин.

Представленные методы воздействия на ПЗП могут привести к необратимым изменениям в скелете горной породы, загрязнениям подземных вод и почво-грунтов.

Работа [3] посвящена комплексному решению проблем гидратообразования в скважинах и промысловых коммуникациях. Наиболее подробно рассматриваются применяемые для борьбы с данным явлением ингибиторы, такие как: метанол, гликоли, хлорид кальция и др. Даны рекомендации по их применению.

Рассмотрено использование забойных и устьевых нагревателей газа, теплоизоляция стволов скважин и шлейфов, экономическая эффективность этих методов. Также приведено применение рассольных пластовых вод для предупреждения гидратообразования в скважинах и шлейфах. Негативное экологическое воздействие на окружающую среду связано с аварийными выбросами и разливами при транспортировке на ГДС ингибиторов, что приводит к загрязнению почв, атмосферы и поверхностных вод.

В [4] рассмотрено влияние на режим эксплуатации газовых скважин деформации пласта в призабойной зоне. Изменение давления и температуры в процессе добычи приводит к ухудшению фильтрационно-емкостных свойств (ФЕС) горных пород, что влияет на технологический режим работы скважин.

Помимо этого, в статье рассмотрены вопросы выбора технологического режима эксплуатации газовых скважин при наличии коррозионно-активных компонентов в составе газа и пластовой воде. Эксперименты показали, что при наличии в газовой смеси углекислого газа увеличение давления среды может привести к росту интенсивности коррозии (при заданной концентрации углекислоты). Особенно важным является минерализация и количество поступающей пластовой воды в скважину. При наличии в пластовой воде органических кислот вместе с углекислым газом в природном газе происходит усиление интенсивности коррозии скважинного и промыслового оборудования.

Одна из самых сложных проблем в эксплуатации газовых скважин связана с присутствием в составе газа сероводорода, что, например, весьма актуально для Астраханского ГКМ. При наличии водного раствора сероводорода большинство сталей при напряженном состоянии быстро разрушаются. Авторами отмечено, что при снижении температуры общая коррозия увеличивается. В результате экспериментальных исследований с образцами различных марок сталей было установлено, что при наличии в газе одновременно сероводорода и углекислого газа разрушение происходило именно из-за действия сероводородной коррозии. Коррозия трубопроводов и оборудования на ГДС может привести к непредвиденным утечкам добываемой смеси и загрязнению окружающей среды.

В работе [5] рассмотрены вопросы рационального использования природных ресурсов и охраны окружающей среды. Авторы отмечают, что качество проектов разработки месторождений существенно влияет на данные показатели. В свою очередь качество проектов зависит от выбранных при проектировании методов и технологий добычи.

Авторами также даны рекомендации по достижению степени экологической чистоты. По мнению авторов, необходимо: создание определенных линейных коммуникаций; компактное размещение скважин; вертикальная компоновка оборудования; создание условий для быстрой локализации аварий.

В работе отмечается, что влияние на ОС оказывают такие факторы, как: расположение скважин; бурение скважин; освоение и исследование скважин, режим их работы; система обвязки скважин с установками комплексной подготовки газа; дожимные компрессорные станции (ДКС), промбаза, установки регенерации ингибиторов. Указанные объекты газодобычи оказывают негативное экологическое воздействие на атмосферу из-за выбросов вредных веществ; на почву из-за сброса отходов производства.

В работе [6] рассмотрены вопросы, связанные с накоплением и выносом жидкости из ствола скважины. Отмечено, что накопленную жидкость возможно удалить при периодической продувке скважин в атмосферу, также через сифонные трубки, остановкой скважин для поглощения жидкости пластом, либо при помощи ввода в скважину пенообразователя.

Рассмотрено также влияние горного давления на параметры разработки газовых и газоконденсатных месторождений. Установлено в результате экспериментальных исследований, что проницаемость и объем пор снижаются из-за деформации пород [6]. Это, в свою очередь, влияет на темп падения давления в пласте. Указанные осложнения могут привести к загрязнению атмосферы и почвогрунтов.

В [7] рассмотрены проблемы эксплуатации скважин в неустойчивых коллекторах. Проблема разрушения ПЗП, особенно вынос песка, на начальных этапах разработки была актуальна еще в самом начале развития нефтегазодобывающей промышленности. Отмечено, что основными факторами, способствующими разрушению ПЗП, являются: неправильная перфорация колонны; создание высоких депрессий на пласт; высокие забойные депрессии; большие дебиты; наличие жидкости в ПЗП.

В свою очередь, механические характеристики и деформации горных пород оказывают значительное влияние на разработку месторождений. Особенно данное влияние проявляется при высоких депрессиях, пластовых давлениях и температурах. Такие условия характерны для месторождений на больших глубинах с аномально высоким пластовым давлением (АВПД). Порода на таких месторождениях оказывается недоуплотненной из-за уравновешивания горного давления пластовым давлением флюида. Отбор при таких условиях происходит при высоких депрессиях, что приводит к разрушению структуры порового пространства и ухудшению коллекторских свойств пласта и снижению показателей разработки. Представленные осложнения могут привести к нерентабельности эксплуатации скважины и перевода ее в статус «ликвидированной», что приводит к серьезным экологическим проблемам для окружающей среды (загрязнение атмосферы, поверхностных и подземных вод).

В работе [8] отмечено, что одним из основных осложняющих факторов является деформация обсадных колонн скважин. Оседание земной поверхности может произойти в процессе разработки месторождения (интервал оседания колеблется в пределах от 10 см до 10 метров и более). Деформация пласта-коллектора происходит в связи с отбором пластового флюида и, как следствие, снижением пластового давления. Зоны с незацементированными участками колонн и в первую очередь участки, находящиеся в зонах кавернообразования, деформируются. Авторами указано, что деформации обсадных колонн можно свести к минимуму при помощи следующих способов: создание зумпфа; проведение периодической инклинометрии ствола скважины; вскрытие продуктивного пласта по нормали к его кровле; тампонаж вязкопластическими жидкостями.

Рассмотрены также проблемы, связанные с коррозией металла. Указано, что почти треть от всей доли металла разрушается из-за коррозии. Факторами, влияющими на интенсивность коррозии, являются: механическое воздействие на металл, структура металла; состав агрессивной смеси; температура и давление среды. Для защиты оборудования от коррозии применяют ингибиторы, которые разделяются на группы: нейтрализаторы, экранирующие ингибиторы, комплексные ингибиторы. Помимо использования ингибиторов на промыслах для защиты оборудования применяют коррозионностойкие металлы, защитные покрытия (металлические, пластмассовые, лакокрасочные, стеклоэмалевые и др.).

Износ оборудования возможен не только по причине наличия агрессивных компонентов, но и механических примесей: частиц песка, барита, гематита, глины, продуктов коррозии. Данные механические частицы могут привести к коррозионно-абразивному износу наземного и подземного оборудования. Отмечено, что одним из наиболее эффективных средств повышения сопротивления стали газоабразивному изнашиванию является термомеханическая обработка. Также причиной разрушения газопроводов является почвенная коррозия и анаэробная биокоррозия. Для защиты трубопроводов от блуждающих токов применяются станции катодной и анодной защиты, протекторная защита и др.

Также авторами рассмотрен вопрос накопления неорганических солей при эксплуатации скважин. Основными факторами отложения солей являются: падение температуры и давления в процессе добычи; смешение пластовых вод с растворами ионного происхождения. Для предупреждения солеобразования применяются дозировочные насосы для подачи ингибитора на устье скважины. Если произошло образование глухой пробки, то необходимо провести капитальный ремонт скважины с подъемом НКТ.

В [9] рассмотрены основные причины снижения производительности скважин. На снижение продуктивности скважин, по мнению автора, влияют следующие геолого-промысловые факторы: снижение ФЕС ласта-коллектора; ухудшение состояния ствола скважины; накопление в стволе скважины жидкости. Отмечено, что на производительность скважин особенное влияние оказывает степень поражения ПЗП, которая зависит от размеров зон кольматации.

Рассмотрен также вопрос изменения продуктивности скважины из-за защемления водяной фазой газонасыщенных областей. Данное явление может привести к серьезному сокращению относительной проницаемости коллектора для углеводородной фазы.

Еще один фактор, влияющий на изменение продуктивности, – выпадение ретроградного конденсата в ПЗП. В связи с резкими изменениями термобарических условий вблизи забоев скважин условия накопления ретроградного конденсата разные. В процессе добычи происходит накопление конденсата в ПЗП, что приводит к уменьшению фазовой проницаемости пласта-коллектора.

Другим фактором снижения ФЕС пласта является выпадение парафинов. Из-за изменения давления и температуры находящиеся в газоконденсатной смеси тяжелые компоненты могут переходить в твердую фазу. Парафины в виде твердой и жидкой фаз могут значительно уменьшить проницаемость и пористость коллектора, что существенно может уменьшить продуктивность газоконденсатных скважин. Осложнения, рассмотренные в данной работе, могут привести к проведению капитальных ремонтных работ либо к ликвидации скважины и загрязнению атмосферы, наземных и подземных вод, почв и почвогрунтов.

Таким образом, в результате анализа имеющихся литературных источников можно сделать вывод, что к негативному экологическому воздействию могут привести как сами осложнения на ГДС, так и мероприятия по их устранению. Все основные осложнения, приведенные в данных источниках, необходимо систематизировать. Например, плановые и внеплановые осложнения можно разделить на осложнения, возникающие под землей и на земле (рис. 2а и 2б).

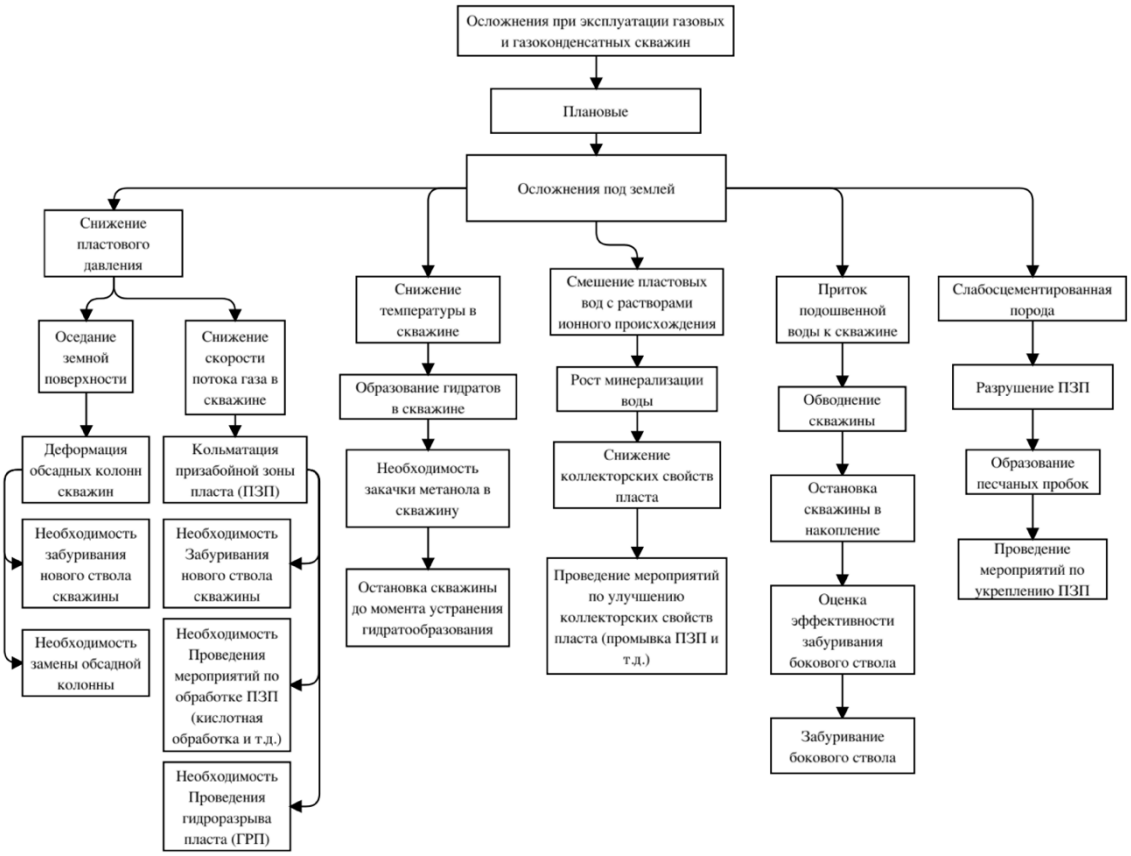


Рисунок 2а – Основные плановые осложнения под землей при эксплуатации скважины

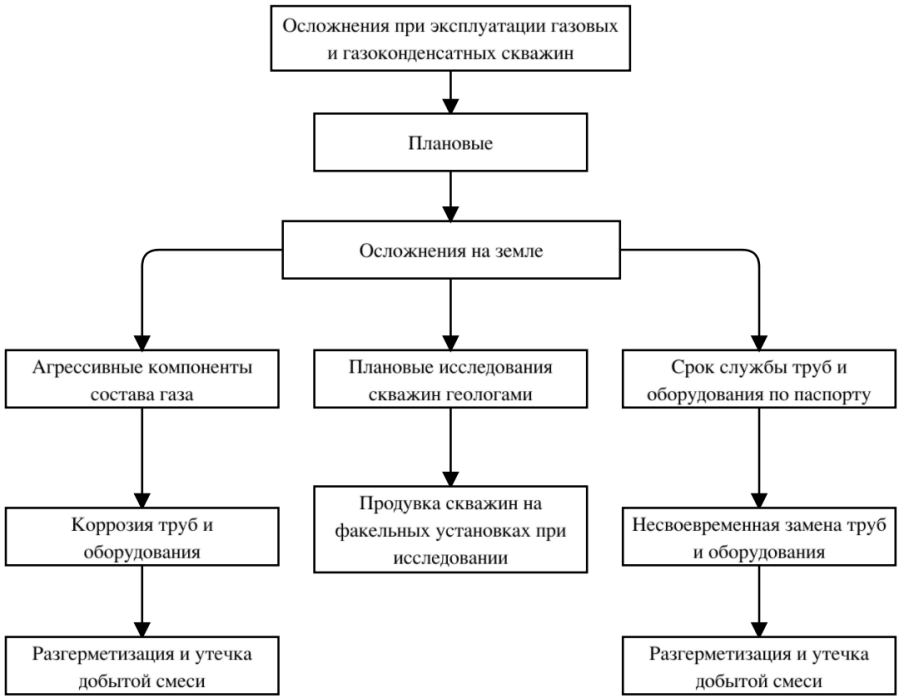


Рисунок 2б – Основные плановые осложнения на земле при эксплуатации скважины

Помимо плановых осложнений на скважине, также могут возникнуть и внеплановые осложнения (рис. 3а и 3б).

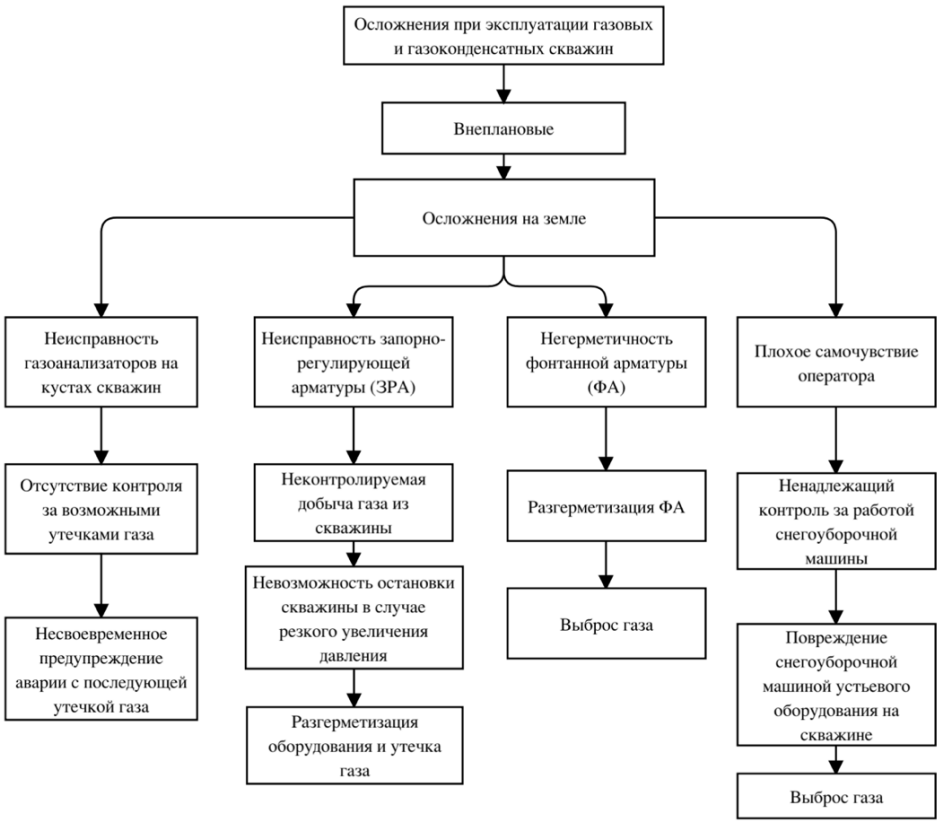


Рисунок 3а – Основные внеплановые осложнения на земле при эксплуатации скважины

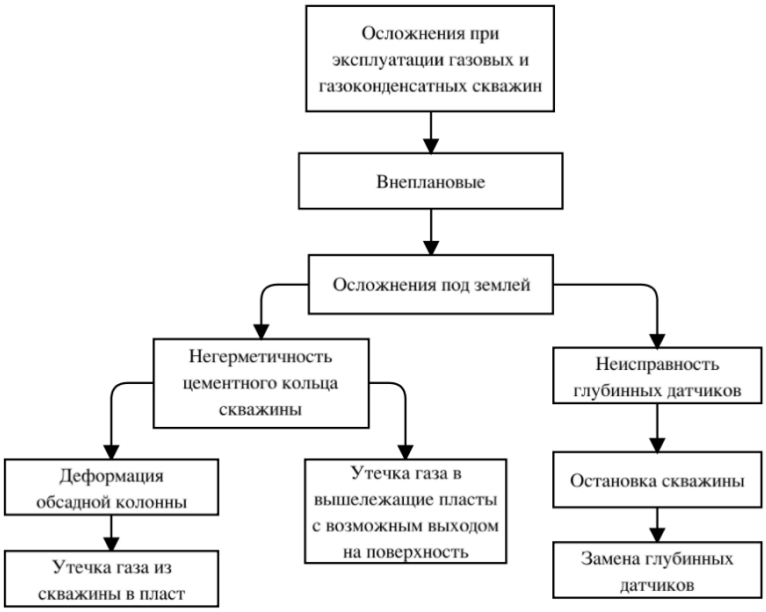


Рисунок 3б – Основные внеплановые осложнения под землей при эксплуатации скважины

Последствия, вызванные осложнениями либо мероприятиями по ликвидации осложнений, приведены в таблице 1.

Таблица 1 – Осложнения при эксплуатации ГДС

Классы осложнений	Условные обозначения	Осложнения	Причины осложнений
ППЗ (плановые под землей)	$ППЗ_1$	Деформация обсадных колонн	Снижение пластового давления
	$ППЗ_2$	Кольматация призабойной зоны	
	$ППЗ_3$	Закачка метанола в скважину	Снижение температуры в скважине
	$ППЗ_4$	Мероприятия по улучшению коллекторских свойств пласта	Смешение пластовых вод с растворами ионного происхождения
	$ППЗ_5$	Забуривание бокового ствола	Приток подошвенной воды
	$ППЗ_6$	Проведение мероприятий по укреплению ПЗП	Слабосцементированная порода
ПНЗ (плановые на земле)	$ПНЗ_1$	Разгерметизация труб и оборудования и утечка добытой смеси	Агрессивные компоненты состава газа
	$ПНЗ_2$	Продувка скважин на факельных установках при исследовании	Плановые исследования скважин геологами
	$ПНЗ_3$	Разгерметизация трубы и оборудования и утечка добытой смеси	Окончание срока службы труб и оборудования по паспорту
ВППЗ (внеплановые под землей)	$ВППЗ_1$	Несвоевременное предупреждение аварии с последующей утечкой газа	Неисправность газоанализаторов на кустах скважин
	$ВППЗ_2$	Разгерметизация оборудования и утечка газа	Неисправность запорно-регулирующей арматуры (ЗРА)
	$ВППЗ_3$	Выброс газа из-за разгерметизации фонтанной арматуры (ФА)	Негерметичность фонтанной арматуры (ФА)
	$ВППЗ_4$	Выброс газа из-за повреждения ФА снегоуборочной машиной	Плохое самочувствие оператора
ВПНЗ (внеплановые на земле)	$ВПНЗ_1$	Утечка газа из скважины в пласт	Негерметичность цементного кольца скважины
	$ВПНЗ_2$	Утечка газа в вышележащие пласты с возможным выходом на поверхность	

Последствия, вызванные осложнениями либо мероприятиями по ликвидации осложнений, в свою очередь, приводят к негативному экологическому воздействию на окружающую среду, что впоследствии может грозить штрафами для компании и затратами на ликвидацию последствий от негативного воздействия на экологию.

В работе [19] предлагалось характеризовать НЭВ на окружающую среду по их длительности (долгосрочные, среднесрочные и краткосрочные), масштабу (региональные, локальные и точечные) и степени устойчивости (постоянные, обратимые и необратимые). Однако необходимо отметить, что НЭВ также характеризуется объемом выбрасываемых в окружающую среду загрязняющих веществ за единицу времени и значимостью каждого воздействия, что, в свою очередь, влияет на их интенсивность. В работе [20] приводится шкала масштабов воздействия и градация экологических последствий, согласно которой можно определить экологическую значимость определенного НЭВ. В соответствии с данной шкалой можно классифицировать: **по масштабу** (выделяют локальные, где площадь воздействия до 1 км^2, ограниченные – площадь до 10 км^2, также местные, где площадь воздействия от 10 до 100 км^2, региональные – площадь более 100 км^2); **по времени** (бывают кратковременные, где время воздействия менее шести месяцев, средней продолжительности – время воздействия от шести месяцев до одного года, продолжительные – от одного до трех лет, многолетние – от трех лет и более); **по интенсивности** (бывают незначительные, где изменения не превышают пределы природной изменчивости; слабое – изменения превышают существующие пределы природной изменчивости, природная среда полностью самовосстанавливается; умеренное – изменения превышают существующие пределы природной изменчивости и приводят к нарушению отдельных компонентов природной среды, но природная среда сохраняет способность к самовосстановлению; сильное – изменения в природной среде приводят к значительным нарушениям компонентов природной среды, отдельные компоненты природной среды теряют способность к самовосстановлению); **по значимости** (низкая – величина воздействия достаточно низка; средняя – величина НЭВ находится в пределах от низкого до уровня, почти нарушающего узаконенный предел; высокая – величина превышает допустимые пределы интенсивности нагрузки на компонент природной среды, либо отмечается воздействие большого масштаба).

Кроме того, ЭП можно разделить на группы по среде, на которую оказывается воздействие (атмосфера, почвы, грунты и т.п.): первая группа «Загрязнение атмосферы» ($ЭП_1$) включает в себя: $ПНЗ_1$, $ПНЗ_2$, $ПНЗ_3$, $ВППЗ_1$, $ВППЗ_2$, $ВППЗ_3$, $ВППЗ_4$, $ВПНЗ_2$; вторая группа «Загрязнение почвы» ($ЭП_2$): $ПНЗ_1$, $ВПНЗ_1$, $ВПНЗ_2$; третья группа «Загрязнение почвогрунтов» ($ЭП_3$): $ППЗ_1$, $ППЗ_2$, $ППЗ_3$, $ППЗ_4$, $ППЗ_5$, $ППЗ_6$; четвертая «Загрязнение подземных вод» ($ЭП_4$): $ППЗ_1$, $ППЗ_2$, $ППЗ_3$, $ППЗ_4$, $ППЗ_5$, $ППЗ_6$; пятая «Загрязнение поверхностных вод» ($ЭП_5$): $ПНЗ_1$, $ВПНЗ_1$.

Все вышеизложенное позволило построить онтологическую схему предметной области, представленную на рисунке 4.

Помимо экологических рисков, возникающих в процессе эксплуатации ГДС, осложнения несут в себе экономические потери (Q) для компаний. Данные осложнения или меры по устранению этих осложнений с определенной вероятностью (P) приводят к экологическим происшествиям (ЭП). ЭП, в свою очередь, характеризуются определенными рисками (R) негативного экологического воздействия на атмосферу, почву, почвогрунты, подземные и поверхностные воды. Риски ЭП определяются вероятностью возникновения ЭП (P), количеством выбросов загрязняющих веществ (V) и суммой ущерба от данных выбросов (S).

Поэтому необходима разработка системы, позволяющей в зависимости от текущего положения на ГДС минимизировать возможные экологические и экономические риски от принимаемых решений по ликвидации осложнений. При этом критерием эффективности должен являться минимальный риск для компании:

$$R \to min, \tag{1}$$

где под R понимается совокупный риск от ЭП.

В свою очередь, R представляет собой сумму рисков, связанных с технологическими осложнениями на ГДС и ЭП, вызванными данными осложнениями:

$$R = R^T + R^{ЭП}, \tag{2}$$

где R^T – риски, связанные с технологическими осложнениями на ГДС; $R^{ЭП}$ – риски, связанные с ЭП на ГДС.

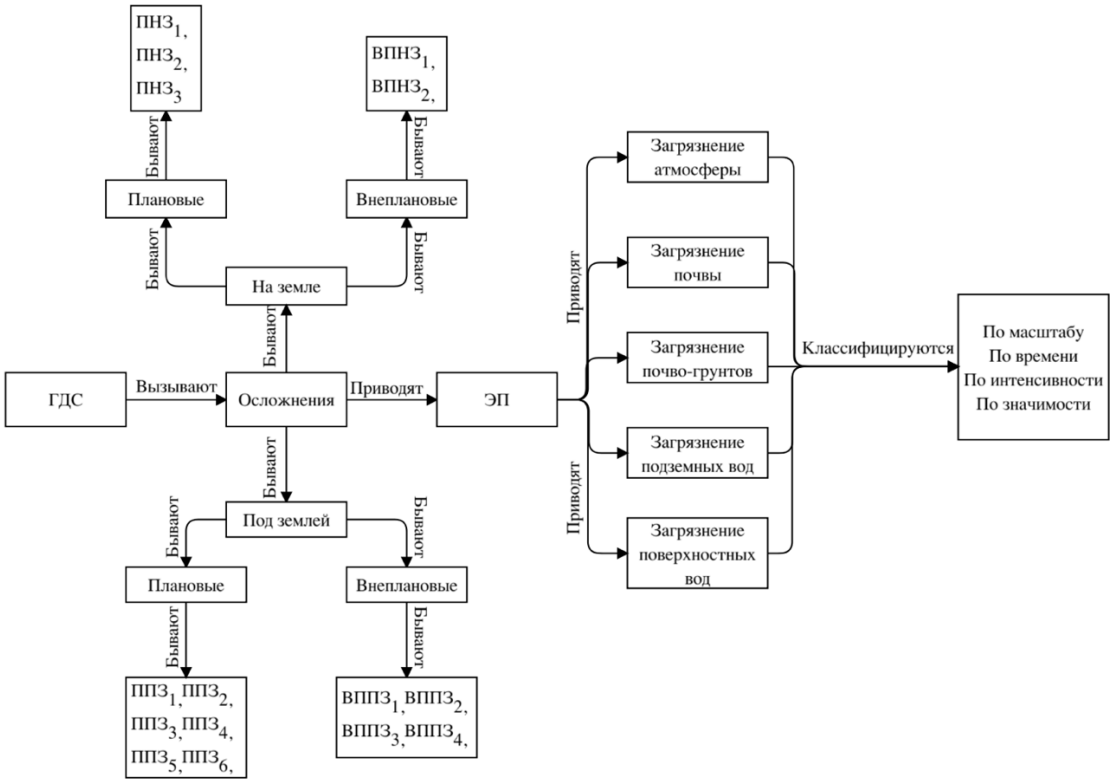


Рисунок 4 – Онтологическая схема загрязнения окружающей среды при эксплуатации ГДС

Такая система поддержки принятия решений (СППР) даст возможность лицу, принимающему решения (ЛПР), более обоснованно выбирать варианты управленческих воздействий с целью минимизации экологических и экономических потерь.

Заключение. Таким образом, проведенный системный анализ технологических осложнений на ГДС, которые могут приводить к негативным экологическим воздействиям на окружающую среду и высоким экономическим рискам для компаний, позволяет перейти к разработке системы, способной в оперативном режиме прогнозировать и предупреждать возможные экологические и экономические риски. В частности, выделенные плановые и внеплановые осложнения под землей и на земле на ГДС и их группировка по среде воздействия на окружающую среду могут быть использованы в формировании Базы знаний (БЗ) для СППР.

Библиографический список

1. Коротаев, Ю. П. Комплексная разведка и разработка газовых месторождений / Ю.П. Коротаев. – Москва : Недра, 1968. – 428 с.

2. Гиматудинов, Ш. К. Разработка и эксплуатация нефтяных, газовых и газоконденсатных месторождений / Ш. К. Гиматудинов, И. И. Дунюшкин, В. М. Зайцев, Ю. П. Коротаев, Е. В. Левыкин, В. А. Сахаров. – Москва : Недра, 1988. – 304 с.

3. Дегтярев, Б. В. Борьба с гидратами при эксплуатации газовых скважин в северных районах / Б. В. Дегтярев, Э. Б. Бухгалтер. – Москва : Недра, 1976. – 198 с.

4. Гриценко, А. И. Руководство по исследованию скважин / А. И. Гриценко, З. С. Алиев, О. М. Ермилов, В. В. Ремизов, Г. А. Зотов – Москва : Наука, 1995. – 523 с.

5. Алиев, З. С. Руководство по проектированию разработки газовых и газонефтяных месторождений г. Печора / З. С. Алиев, В. В. Бондаренко. – Печора : Печорское время, 2002. – 895 с.

6. Мирзаджанзаде, А. Х. Основы технологии добычи газа / А. Х. Мирзаджанзаде, О. Л. Кузнецов, К. С. Басниев, З. С. Алиев. – Москва : Недра, 2003. – 880 с.

7. Зотов, Г. А. Эксплуатация скважин в неустойчивых коллекторах / Г. А. Зотов, А. В. Динков, В. А. Черных. – Москва : Недра, 1987. – 172 с.

8. Вяхирев, Р. И. Теория и опыт добычи газа / Р. И. Вяхирев, Ю. П. Коротаев, Н. И. Кабанов. – Москва : ОАО «Издательство «Недра», 1998. – 479 с.

9. Тер-Саркисов, Р. М. Разработка месторождений природных газов / Р. М. Тер-Саркисов. – Москва : ОАО «Издательство «Недра», 1999. – 659 с.

10. Информационно-технический справочник по наилучшим доступным технологиям. – Москва : Бюро НТД, 2017. – 271 с.

11. Безродный, Ю. Г. Разработка методов обеспечения охраны окружающей среды при проектировании и строительстве нефтегазовых скважин : дис. ... д-ра техн. наук: 25.00.15 / Юрий Георгиевич Безродный. – Москва, 2009. – 369 с.

12. Макаренкова, И. Ю. Экологическая оценка воздействия нефтегазодобывающей деятельности на водные объекты Среднего Приобья : автореф. дис. … канд. геогр. наук / И. Ю. Макаренкова. – Ростов н/Д, 2007. – 25 с.

13. Булатов, А. И. Актуальные проблемы охраны окружающей среды при бурении скважин / А. И. Булатов, В. И. Рябчепко, В. Ю. Шеметов // Нефтяное хозяйство. – 1988. – № 6. – С. 5–9.

14. Об охране окружающей среды : федеральный закон от 10 января 2002 года № 7-ФЗ ; в ред. от 30.12.21 // Собрание законодательства РФ. – 2021. – № 52. – Ст. 1704.

15. О Стратегии экологической безопасности Российской Федерации на период до 2025 года : указ Президента Российской Федерации [издан Президентом 19 апреля 2017 г. № 176] // Собрание законодательства РФ. – 2017. – № 17. – Ст. 2546.

16. Агадулин, И. И. Экологические аспекты негерметичности заколонного пространства в скважинах различного назначения / И. И. Агадулин, В. Н. Игнатьев, Р. Ю. Сухоруков // Нефтегазовое дело. – 2011. – № 4. – С. 82–90.

17. Зорина, О. Я. Система удаленного газомониторинга приустьевого пространства ликвидированных глубоких скважин / О. Я. Зорина, В. В. Кудинов // Защита окружающей среды в нефтегазовом комплексе. – 2012. – № 8. – С. 26–28.

18. Султанов, В. Г. Проблемы качественного крепления нефтяных скважин при их строительстве и обеспечение в последующем их надежной ликвидации, консервации / В. Г. Султанов, Л. В. Примак // Механизация строительства. – 2014. – № 7. – С. 44–48.

19. ИТС 29-2017 Добыча природного газа, Федеральное агентство по техническому регулированию и метрологии.

20. Предварительная оценка воздействия на окружающую среду к проекту разработки месторождения Айрантыкар, Филиал ТОО «КМГ ИНЖИНИРИНГ» «Казахский научно-исследовательский и проектный институт нефти и газа», г. Актау. – С. 119–127.

21. Предеин, А. П. Осложнения и аварии при строительстве нефтяных и га зовых скважин : учеб. пособие / А. П. Предеин. – Пермь : Изд-во Перм. нац. исслед. политехн. ун-та, 2014. – 381 с.

22. Компанию «Норникеля» оштрафовали на рекордные 146 млрд рублей за разлив нефти // BBC NEWS. – Режим доступа: https://www.bbc.com/russian/news-55950310, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 12.01.2022).

23. В Оренбургской области на магистральном газопроводе произошел взрыв // ТАСС. – Режим доступа: https://tass.ru/proisshestviya/10763889, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения 10.01.2022).

24. Sun, Y. Q. A Critical Review of Risks, Characteristics, and Treatment Strategies for Potentially Toxic Elements in Wastewater from Shale Gas Extraction / Y. Q. Sun, D. Wang, D. C. W. Tsang et al. // Environment International. – 2019. – Vol. 125. – P. 452–469.

25. Monika, Wójcik. Environmental Risk Assessment for Exploration and Extraction Processes of Unconventional Hydrocarbon Deposits of Shale Gas and Tight Gas: Pomeranian and Carpathian Region Case Study as Largest Onshore Oilfields / Monika Wójcik, Wojciech Kostowski // Journal of Earth Science. – 2020. – № 31 (1). – P. 215–222.

26. Mac Kinnon, M. A. The role of natural gas and its infrastructure in mitigating greenhouse gas emissions, improving regional air quality, and renewable resource integration / M. A. Mac Kinnon, J. Brouwer, S. Samuelsen // Prog. Energy Combust. Sci. – 2018. – № 64. – P. 62–92.

27. Mackay, D. J. C. Potential greenhouse gas emissions associated with shale gas extraction and use / D. J. C. Mackay, T. J. Stone. – London, United Kingdom : Department of Energy and Climate Change. – 2013. – № 50.

28. McKenzie, L. M. Human health risk assessment of air emissions from development of unconventional natural gas resources / L. M. McKenzie, R. Z. Witter, L. S. Newman, J. L. Adgate // Sci. Total Environ. – 2012. – Vol. 424. – P. 79–87.

29. Shonkoff, S. B. C. Environmental public health dimensions of shale and tight gas development / S. B. C. Shonkoff, J. Hays, M. L. Finkel // Environ. Health Perspect. – 2014. – Vol. 122 (8). – P. 787–795.

30. Annika, W. Walters. Multiple approaches to surface water quality assessment provide insight for small streams experiencing oil and natural gas development / Annika W. Walters, Carlin E. Girard, Richard H. Walker, Aïda M. Farag, David A. Alvarez // Integrated Environmental Assessment and Management. – 2019. – № 15 (3). – P. 385–397.

References

1. Korotaev, Yu. P. *Kompleksnaya razvedka i razrabotka gazovykh mestorozhdeniy* [Integrated exploration and development of gas fields]. Moscow, Nedra Publ., 1968. 428 p.

2. Gimatudinov, Sh. K. Dunyushkin, I. I., Zaytsev, V. M., Korotaev, Yu. P., Levykin, E. V., Sakharov, V. A. *Razrabotka i ekspluatatsiya neftyanykh, gazovykh i gazokondensatnykh mestorozhdeniy* [Development and operation of oil, gas and gas condensate fields]. Moscow, Nedra Publ., 1988. 304 p.

3. Degtyarev, B. V., Bukhgalter, E. B. *Borba s gidratami pri ekspluatatsii gazovykh skvazhin v severnykh rayonakh* [Fight against hydrates during the operation of gas wells in the northern regions]. Moscow, Nedra Publ., 1976. 198 p.

4. Gricenko, A. I., Aliev, Z. S., Ermilov, O. M., Remizov, V. V., Zotov, G. A. *Rukovodstvo po issledovaniyu skvazhin* [Guidelines for the study of wells]. Moscow, Nauka Publ., 1995. 523 p.

5. Aliev, Z. S., Bondarenko, V. V. *Rukovodstvo po proektirovaniyu razrabotki gazovykh i gazoneftyanykh mestorozhdeniy goroda Pechora* [Guidelines for the design of the development of gas and gas-oil fields in the city of Pechora]. Pechora, Pechorskoe vremya Publ., 2002. 895 p.

6. Mirzadzhanzade, A. H., Kuznetsov, O. L., Basniev, K. S., Aliev, Z. S. *Osnovy tekhnologii dobychi gaza* [Fundamentals of gas production technology]. Moscow, Nedra Publ., 2003. 880 p.

7. Zotov, G. A., Dinkov, A. V. Chernykh, V. A. *Ekspluatatsiya skvazhin v neustoychivykh kollektorakh* [Operation of wells in unstable reservoirs]. Moscow, Nedra Publ., 1987. 172 p.

8. Vyakhirev, R. I., Korotaev, Yu. P., Kabanov, N. I. *Teoriya i opyt dobychi gaza* [Theory and experience of gas production]. Moscow, Publishing House "Nedra" OJSC, 1998. 479 p.

9. Ter-Sarkisov, R. M. *Razrabotka mestorozhdeniy prirodnykh gazov* [Development of natural gas deposits]. Moscow, Publishing House "Nedra" OJSC, 1999. 659 p.

10. *Informatsionno-tekhnicheskiy spravochnik po nailuchshim dostupnym tekhnologiyam* [Information and technical guide to the best available technologies]. Moscow, BAT Bureau, 2017. 271 p.

11. Bezrodnyy, Yu. G. *Razrabotka metodov obespecheniya okhrany okruzhayushchey sredy pri proektirovanii I stroitelstve neftegazovykh skvazhin : dissertatsiya ... doktora tekhnicheskikh nauk: 25.00.15* [Development of methods for ensuring environmental protection in the design and construction of oil and gas wells : dissertation ... of Doct. of Tech. Sciences: 25.00.15]. Moscow, 2009. 369 p.

12. Makarenkova, I. Yu. *Ekologicheskaya otsenka vozdeystviya neftegazodobyvayushchey deyatelnosti na vodnye obekty Srednego Priobya : avtoreferat dissertatsii ... kandidata geograficheskikh nauk* [Ecological assessment of the impact of oil and gas production on water bodies of the Middle Ob : abstract of dissertation … of Cand. Of Geogr. Sciences]. Rostov on Don, 2007. 25 p.

13. Bulatov, A. I., Ryabchepko, V. I., Shemetov, V. Yu. Aktualnye problemy okhrany okruzhayushchey sredy pri burenii skvazhin [Actual problems of environmental protection during well drilling]. *Neftyanoe khozyaystvo* [Oil industry], 1988, no. 6, pp. 5–9.

14. Ob okhrane okruzhayushchey sredy : federalnyy zakon ot 10 yanvarya 2002 goda № 7-FZ; v redaktsii ot 30.12.21 [On environmental protection : federal law of January 10, 2002 No. 7-FZ ; in ed. of 12.30.21]. *Sobranie zakonodatelstva RF* [Collection of legislation of the Russian Federation], 2021, no. 52, art. 1704.

15. O Strategii ekologicheskoy bezopasnosti Rossiyskoy Federatsii na period do 2025 goda: ukaz Prezidenta Rossiyskoy Federatsii [izdan Prezidentom 19 aprelya 2017 goda № 176] [On the Strategy for Environmental Security of the Russian Federation for the period up to 2025: Decree of the President of the Russian Federation [issued by the President on April 19, 2017 No. 176]]. *Sobranie zakonodatelstva RF* [Collection of legislation of the Russian Federation], 2017, no. 17, art. 2546.

16. Agadulin, I. I., Ignatev, V. N., Sukhorukov, R. Yu. Ekologicheskie aspekty negermetichnosti zakolonnogo prostranstva v skvazhinah razlichnogo naznacheniya [Remote gas monitoring system near the wellhead of abandoned deep wells]. *Neftegazovoe delo* [Oil and Gas Business], 2011, no. 4, pp. 82–90.

17. Zorina, O. Ya., Kudinov, V. V. Sistema udalennogo gazomonitoringa priustevogo prostranstva likvidirovannyh glubokih skvazhin [Remote gas monitoring system near the wellhead of abandoned deep wells]. *Zashchita okruzhayushchey sredy v neftegazovom komplekse* [Environmental protection in the oil and gas complex], 2012, no. 8, p. 26–28.

18. Sultanov, V. G., Primak, L. V. Problemy kachestvennogo krepleniya neftyanykh skvazhin pri ikh stroitelstve i obespechenie v posleduyushchem ikh nadezhnoy likvidatsii, konservatsii [Problems of high-quality casing of oil wells during their construction and ensuring their subsequent reliable liquidation, conservation]. *Mekhanizatsiya stroitelstva* [Construction Mechanization], 2014, no. 7, pp. 44–48.

19. *ITS 29-2017 Dobycha prirodnogo gaza, Federalnoe agentstvo po tekhnicheskomu regulirovaniyu I metrologii* [Natural gas production, Federal Agency for Technical Regulation and Metrology].

20. *Predvaritelnaya otsenka vozdeystviya na okruzhayushchuyu sredu k proektu razrabotki mestorozhdeniya Ayrantykar, Filial TOO «KMG INZHINIRING» «Kazahskiy nauchno-issledovatelskiy i proektnyy institut nefti i gaza»* [Preliminary environmental impact assessment for the Ayrantikar field development project, Branch of LLP "KMG ENGINEERING" "Kazakh Research and Design Institute of Oil and Gas"]. Aktau, pp. 119–127.

21. Predein, A. P. *Oslozhneniya i avarii pri stroitelstve neftyanykh i gazovykh skvazhin : uchebnoe posobie* [Complications and accidents in the construction of oil and gas wells : textbook]. Perm, Publishing House of Perm National Research Polytechnical, 2014. 381 p.

22. Kompaniyu "Nornikelya" oshtrafovali na rekordnye 146 milliardov rubley za razliv nefti [Norilsk Nickel fined a record 146 billion rubles for an oil spill]. *BBC NEWS*. Available at: https://www.bbc.com/russian/news-55950310

23. *V Orenburgskoy oblasti na magistralnom gazoprovode proizoshel vzryv* [In the Orenburg region, an explosion occurred on the main gas pipeline]. *TASS* [TASS]. Available at: https://tass.ru/proisshestviya/10763889.

24. Sun, Y. Q., Wang, D., Tsang, D. C. W., et al. A Critical Review of Risks, Characteristics, and Treatment Strategies for Potentially Toxic Elements in Wastewater from Shale Gas Extraction. *Environment International*, 2019, vol. 125, pp. 452–469.

25. Monika, Wójcik, Wojciech, Kostowski. Environmental Risk Assessment for Exploration and Extraction Processes of Unconventional Hydrocarbon Deposits of Shale Gas and Tight Gas: Pomeranian and Carpathian Region Case Study as Largest Onshore Oilfields. *Journal of Earth Science*, 2020, no. 31 (1), pp. 215–222.

26. Mac Kinnon, M. A., Brouwer, J., Samuelsen, S. The role of natural gas and its infrastructure in mitigating greenhouse gas emissions, improving regional air quality, and renewable resource integration. *Prog. Energy Combust. Sci.*, 2018, no. 64, pp. 62–92.

27. Mackay, D. J. C., Stone, T. J. *Potential greenhouse gas emissions associated with shale gas extraction and use*. London, United Kingdom: Department of Energy and Climate Change, 2013, no. 50.

28. McKenzie, L. M., Witter, R. Z., Newman, L. S., Adgate, J. L. Human health risk assessment of air emissions from development of unconventional natural gas resources. *Sci. Total Environ.*, 2012, vol. 424, pp. 79–87.

29. Shonkoff, S. B. C., Hays, J., Finkel, M. L. Environmental public health dimensions of shale and tight gas development. *Environ. Health Perspect.*, 2014, vol. 122 (8), pp. 787–795.

30. Annika, W. Walters, Carlin, E. Girard, Richard, H. Walker, Aïda, M. Farag, David, A. Alvarez. Multiple approaches to surface water quality assessment provide insight for small streams experiencing oil and natural gas development. *Integrated Environmental Assessment and Management*, 2019, no. 15 (3), pp. 385–397.

УПРАВЛЕНИЕ В ОРГАНИЗАЦИОННЫХ СИСТЕМАХ

УДК 004.021

МЕТОД АНАЛИЗА ТЕНДЕНЦИЙ РАЗВИТИЯ ТЕХНОЛОГИЙ УПРАВЛЕНИЯ ЭФФЕКТИВНОСТЬЮ АКТИВОВ[1]

Статья поступила в редакцию 01.02.2022, в окончательном варианте – 18.02.2022.

Нгуен Тхань Вьет, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28; Фам Ван Донг университет, Куанг Нгай, Вьетнам,
аспирант, ORCID: 0000-0003-3805-5958, e-mail: vietqn1987@gmail.com
Кравец Алла Григорьевна, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28; Государственный университет «Дубна», 141982, Российская Федерация, Московская область, г. Дубна, ул. Университетская, 19,
доктор технических наук, профессор, ORCID: 0000-0003-1675-8652, e-mail: agk@gde.ru
Щербаков Максим Владимирович, заведующий кафедрой САПРиПК, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28,
доктор технических наук, профессор, ORCID: 0000-0001-7173-4499, e-mail: maxim.shcherbakov@gmail.com

Чтобы поддерживать конкурентоспособность в среде быстро меняющейся науки, важно отслеживать развитие существующих технологий, открывать новые и многообещающие технические решения. Предприятиям необходимо разрабатывать стратегию инновационного развития на основе прогноза новых технологий, чтобы получить конкурентное преимущество при использовании ограниченных ресурсов. В работе предложен метод анализа тенденции развития технологий в области управления эффективностью активов (Asset Performance Management – APM), определены самые влиятельные факторы (изобретатели, ученые, организации). В качестве источника данных использована система The Lens – флагманский проект общественной организации Cambia. В исходный датасет вошли 115677 патентов, после предобработки которых с учетом самой ранней даты приоритета в периоде 2010–2021 гг. и группировки по простым семействам получено 30706 патентов. Из 18073 научных публикаций, посвященных теме области APM, отобраны 9069 в периоде 2010–2021 гг. для дальнейшего анализа. Кроме того, были исследованы 29943 научные публикации, цитируемые в найденных патентах по APM. В результате определены основные направления исследований и их тенденции в области APM за последние 12 лет (2010–2021), выявлены топ-20 тематик исследования и 9 тематик восходящего тренда в публикациях.

Ключевые слова: управление эффективностью активов, анализ данных, анализ технологии, патент, научные публикации, Asset Performance Management (APM)

METHOD FOR DEVELOPMENT TRENDS ANALYSIS OF ASSET PERFORMANCE MANAGEMENT TECHNOLOGIES

The article was received by the editorial board on 01.02.2022, in the final version – 18.02.2022.

Nguyen Thanh Viet, Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation; Pham Van Dong University, Quang Ngai, Vietnam,
post-graduate student, e-mail: vietqn1987@gmail.com
Kravets Alla G., Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation,
Doct. Sci. (Engineering), Professor, ORCID: 0000-0003-1675-8652, e-mail: agk@gde.ru
Shcherbakov Maksim V., Head of the CAD&SD Department, Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation,
Doct. Sci. (Engineering), Professor, ORCID: 0000-0001-7173-4499, e-mail: maxim.shcherbakov@gmail.com

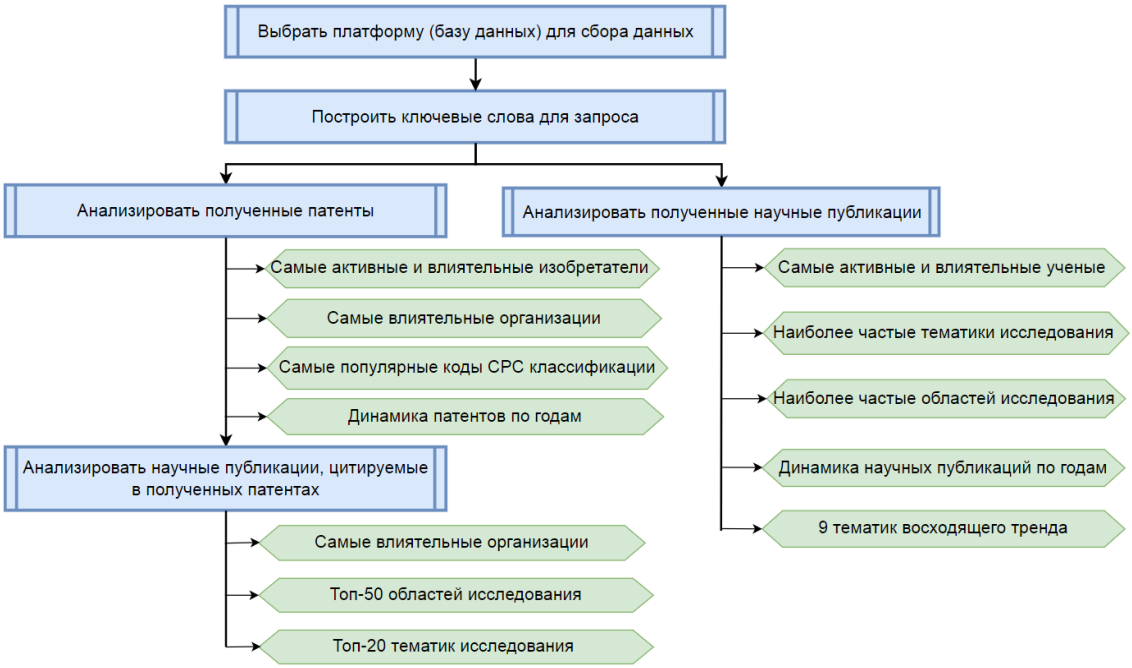
In order to remain competitive in a rapidly changing science environment, it is important to keep track of the development of existing technologies, to discover new and promising technical solutions. An enterprise needs to develop an innovative development strategy based on a forecast of new technologies in order to gain a competitive advantage while using limited resources. The paper proposes a method for analyzing the trend of technology development in the field of asset performance management (APM), implemented the selection of the most influential factors (inventors,

[1]Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90092.

scientists, organizations). The Lens system, the flagship project of the social organization Cambia, was used as a data source. The initial dataset included 115677 patents, after pre-processing, taking into account the earliest priority date in the period 2010-2021 and grouping by simple families, 30706 patents were received. Out of 18073 scientific publications devoted to the topic of the APM field, 9069 were selected in the period 2010–2021 for further analysis. In addition, 29943 scientific publications cited in found APM patents were also analyzed. The main areas of research and their trends in the field of APM over the past 12 years (2010–2021) were explored. Moreover, the top-20 research subjects and 9 uptrend subjects in publications were also identified.

Keywords: asset performance management, data analysis, technology analysis, patent, scientific publications, Asset Performance Management (APM)

Graphical annotation (Графическая аннотация)



Введение. Новые технологии способны радикально изменить ситуацию как на мировых рынках, так и на отдельном предприятии. Речь идет о таких вещах, как облачные решения, управление большими данными, промышленный интернет вещей, комплексное системное моделирование и аналитика. По отдельности они обеспечивают предприятиям возможность стратегически планировать, прогнозировать и оптимизировать операции. Вместе же эти технологии составляют мощный инструментарий, который позволяет снижать издержки и увеличивать производительность [1].

С другой стороны, в процессе управления промышленным предприятием одной из актуальных проблем становится анализ эффективности управления его активами. В условиях рыночной экономики активы являются дорогостоящим товаром, который формирует основную часть рыночной стоимости предприятия как имущественного комплекса. В настоящее время самый распространенный подход к обеспечению непрерывной эксплуатации оборудования – это профилактическое обслуживание. Оно проводится через регулярные промежутки времени, чтобы уменьшить вероятность отказов. Однако в большинстве случаев такая практика приводит либо к излишнему, либо к недостаточному обслуживанию из-за непредсказуемых ситуаций, различий в возрасте оборудования и условиях эксплуатации.

Управление эффективностью активов (Asset Performance Management – APM) – это подход к управлению активами, который определяет приоритеты бизнес-целей в дополнение к традиционным целям надежности и доступности активов. APM стал основным инструментом цифровой трансформации промышленных компаний. Цифровая трансформация позволяет компаниям сменить стратегию техобслуживания с реактивной на проактивную [2]. Эта стратегия позволяет персоналу заблаговременно предотвращать дорогостоящие отказы и максимально эффективно использовать активы в течение всего их жизненного цикла [3]. Современный APM сочетает в себе традиционные методы управления активами с новыми цифровыми технологиями для повышения надежности, выполнения технического обслуживания и обеспечения эффективности бизнеса [4].

В работе [5] представлен современный подход к повышению эффективности управления техническим обслуживанием и ремонтами (ТОиР) производственных активов предприятия с помощью информационных центров управления производственными активами на базе SAP ERP (ERP – система управления ресурсами предприятия). Решение позволяет повысить производительность труда специалистов ТОиР за счет концентрации всех действий по планированию и управлению процессами ТОиР в «едином информационном окне».

Аналогично в исследовании [6] приведен поиск возможных путей создания единой цифровой модели промышленного предприятия для повышения эффективности управления его непроизводственными активами. При этом в статье проведен подробный анализ различных информационных систем, применяемых на всех стадиях жизненного цикла непроизводственных активов промышленного предприятия, начиная с прединвестиционных исследований и заканчивая стадиями эксплуатации и ликвидации. Авторами предложена единая информационная модель активов предприятия, основанная на интеграции ERP-системы предприятия и BIM-моделей его зданий и сооружений.

Тем самым средства автоматизации объединяют операции и позволяют экспертам в реальном времени видеть, что происходит внутри бизнес-процессов. Мобильная связь обеспечивает доступ к данным отовсюду и улучшает сотрудничество бизнес-подразделений. Машинное обучение, искусственный интеллект и облачные технологии также помогают реализовать прогнозные стратегии, которые существенно сокращают незапланированные простои (рис. 1). Подход, ориентированный на данные, повышает эффективность актива, что является дополнительным экономическим преимуществом для эксплуатирующего предприятия [7, 8].

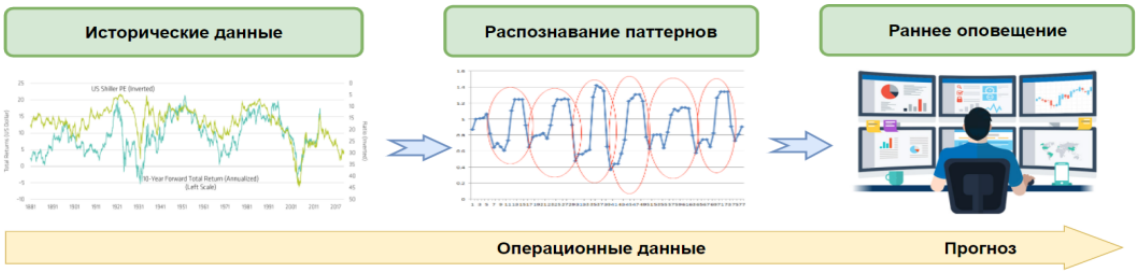


Рисунок 1 – Предиктивная аналитика актива

Помимо этого, работа [9] рассматривает движущие силы бизнеса (business drivers), проблемы и инновации в отрасли, направленные на повышение надежности энергосистемы с помощью АРМ-технологий. В ней подчеркивается роль оцифровки в принятии надежных и основанных на данных решений путем объединения информации из разных организационных хранилищ и проведения методичного и последовательного анализа. Также предложено использовать цифровые технологии, такие как мобильные инструменты, промышленный интернет, большие данные и предиктивный анализ. На реальных примерах показано, как реализуется проект АРМ и как заказчик за короткое время добился чрезвычайно высокой окупаемости инвестиций с помощью практических технологий.

Таким образом, целью данной работы является разработка и реализация метода анализа современных достижений, значительных направлений исследований и тенденций развития технологий как в предметной области АРМ, так и в смежных областях. Практическая значимость данного метода анализа состоит в выявлении сильных и слабых сторон технологической деятельности в выбранной области и позволяет обозначить возможности увеличения конкурентных преимуществ за счет надлежащего использования технологий, а также определить доступные технологии для улучшения своих продуктов и процессов.

1. Предложенный метод анализа предметной области АРМ. Общий вид предложенного метода анализа предметной области АРМ представлен на рисунке 2.

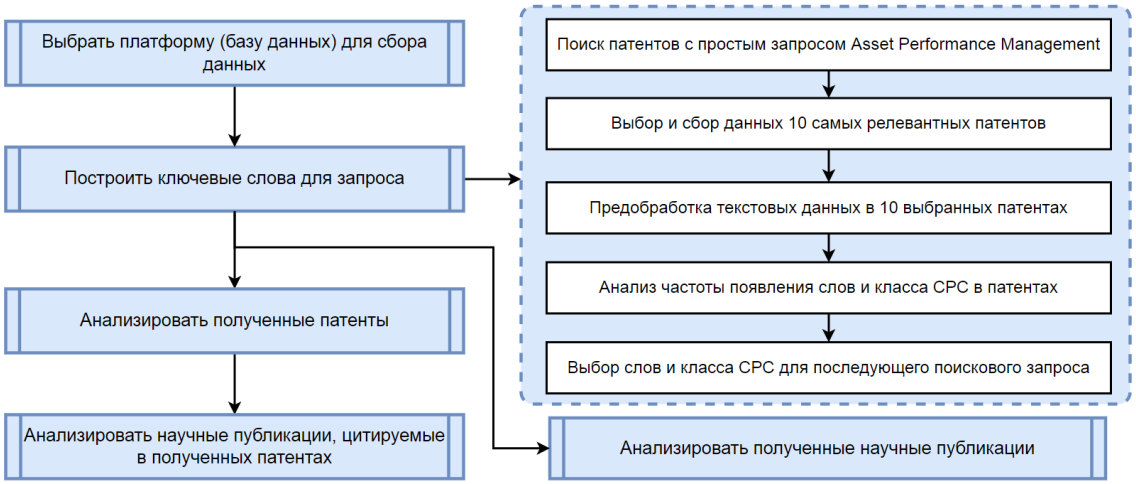


Рисунок 2 – Общий метод для анализа (CPC – Cooperative Patent Classification)

1.1. Выбор платформы – источника патентных документов для анализа. Решение о том, какой источник патентной информации будет использоваться, является важным на начальном этапе исследовательской работы. Стоимость сбора данных для анализа следует сопоставить со временем, которое потребуется на работу с данными, и их полнотой.

Хотя только сами государственные патентные органы генерируют авторитетные патентные данные (первичные источники), существует ряд различных вторичных источников патентных документов и информации, которые обычно используются для создания отчетов о патентовании. По своей природе вторичные источники обычно включают информацию о семействах патентных документов. В некоторых базах данных эта информация о семействе используется для сокращения семейств в списке результатов, то есть список результатов поиска будет включать только один документ из семейства патентов.

Несколько патентных ведомств поддерживают вторичные патентные базы данных, которые позволяют искать патенты из нескольких стран вместе. Например, это PATENTSCOPE от WIPO, Esp@cenet от Европейского патентного ведомства или DEPATISNET от Немецкого ведомства по патентам и товарным знакам.

Есть еще несколько бесплатных онлайн-сервисов по поиску патентов. Их характеристики аналогичны предложениям патентных ведомств, хотя их охват по странам может быть меньше. Иногда они предлагают преимущества по сравнению с сайтами патентных ведомств, поскольку их пользовательский интерфейс зачастую немного более отлажен и удобен для конечного пользователя. Иногда они также предлагают дополнительные функции, которых обычно нет на сайтах патентных ведомств. Например, The Lens от Cambia и Patent Inspiration от CREAX предлагают некоторые функции статистического анализа и визуализации.

Среди них Cambia – The Lens (http://www.lens.org) – всемирный открытый бесплатный полнотекстовый ресурс патентной информации. The Lens, флагманский проект общественной организации Cambia, направлен на поиск, объединение и связывание различных наборов открытых знаний, включая научные работы и патенты, для информирования об открытии, анализе, принятии решений и партнерстве с ориентированным на человека пользовательским интерфейсом, основанном на открытой веб-платформе Lens.org с инструментами для решения различных задач анализа.

В течение более 20 лет развития, благодаря поддержке известными благотворительными организациями, The Lens собирает, очищает, объединяет, нормализует и обслуживает более 225 миллионов научных работ, 127 миллионов глобальных патентных записей и более 370 миллионов патентных последовательностей с обширными метаданными, взятыми из различных источников данных [10].

Таким образом, в данной работе платформа The Lens использована для сбора данных и выполнения анализа тенденций в предметной области APM.

1.2. Построение набора ключевых слов для поискового запроса. Для начала был реализован поиск патентов на платформе The Lens с простым поисковым запросом «Asset Performance Management». В результате был получен набор 95032 патентов. Далее авторы изучили небольшое количество патентов и выбрали 10 самых релевантных документов, из которых данные: заголовок, аннотация, классификация (CPC) – были извлечены для последующего анализа. Предварительная

обработка данных включает в себя удаление знаков препинания и стоп-слов из стандартного списка, а также лемматизацию с помощью библиотеки NLP spaCy. Новое свойство Text образуется путем комбинирования заголовков и аннотаций, а затем разбивается на лексемы.

В результате получен график частоты появления в патентах топ-30 слов, как показано на рисунке 3. На основе этого следующие слова были включены в поисковый запрос: data, model, digital, predict, analy* (analyze, analytic, analysis).

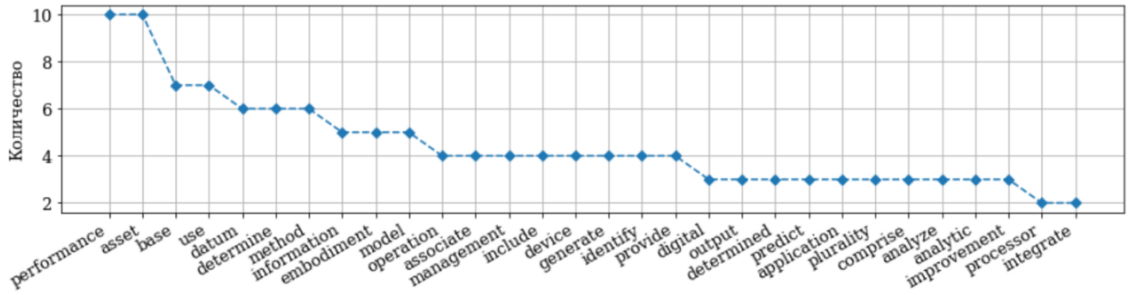


Рисунок 3 – Частоты появления топ-30 слов в 10 релевантных патентах

Кроме того, была осуществлена статистическая обработка 4 первых уровней классов совместной патентной классификации Европейского патентного ведомства и Ведомства по патентам и товарным знакам США – Cooperative Patent Classification (CPC), результаты которой представлены в рисунке 4.

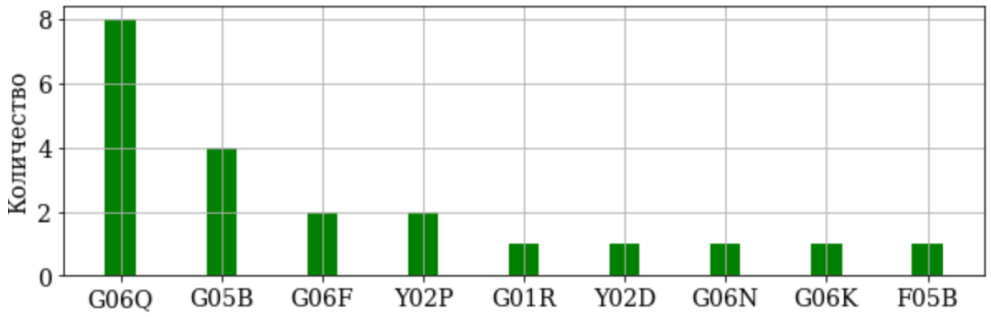


Рисунок 4 – Частоты появления CPC классов в 10 релевантных патентах

На основе этого в выборку были включены классы:

- G06Q (системы или методы обработки данных, специально адаптированные для административных, коммерческих, финансовых, управленческих, надзорных или прогнозных целей);
- G05B (системы контроля или регулирования в целом; функциональные элементы таких систем; меры по мониторингу или тестированию таких систем или элементов);
- G06F (компьютерные системы электронной цифровой обработки данных на основе конкретных вычислительных моделей).

Таким образом, составлен следующий поисковый запрос:

```
((asset AND performance AND manage* AND (data || model || digital || predict || analy*)) || ("asset performance management")) AND (class_cpc.symbol:G06Q* || G05B* || G06F *).
```

Данный запрос подается в поисковую систему платформы The Lens и получает в результате набор 115677 патентов. Далее выполнена фильтрация патентов с самой ранней датой приоритета в периоде 2010–2021 гг. и группировка по простым семействам (Simple Families). В результате получено 30706 патентов, как показано на рисунке 5.

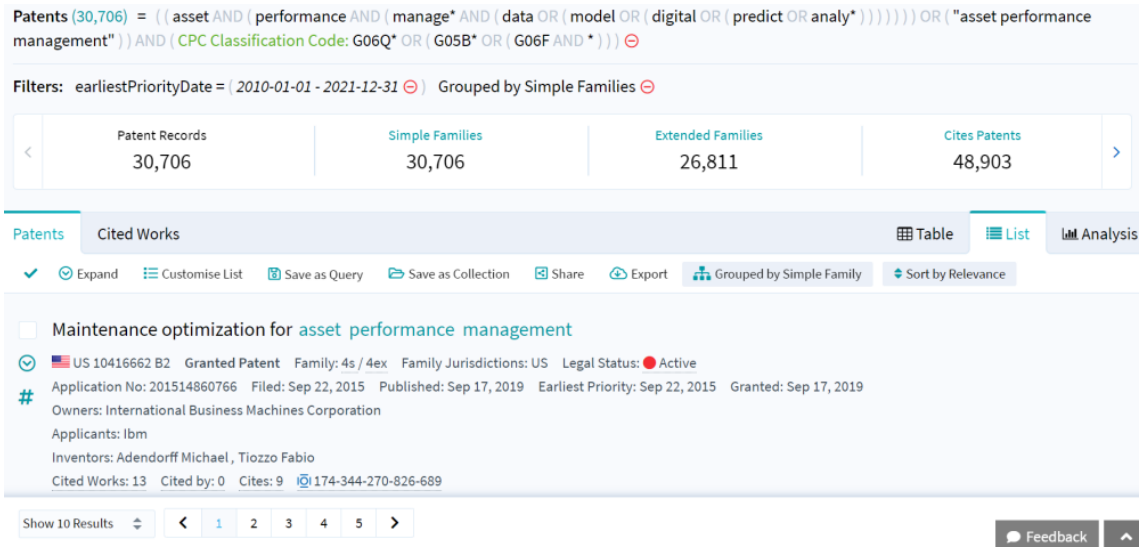


Рисунок 5 – Результат поиска и фильтрования патентов по составленному запросу

Для поиска научных публикаций приведен подобный поисковый запрос (для заголовка или аннотации):

```
((asset AND performance AND manage* AND (data || model || digital || predict || analy*)) || ("asset performance management")).
```

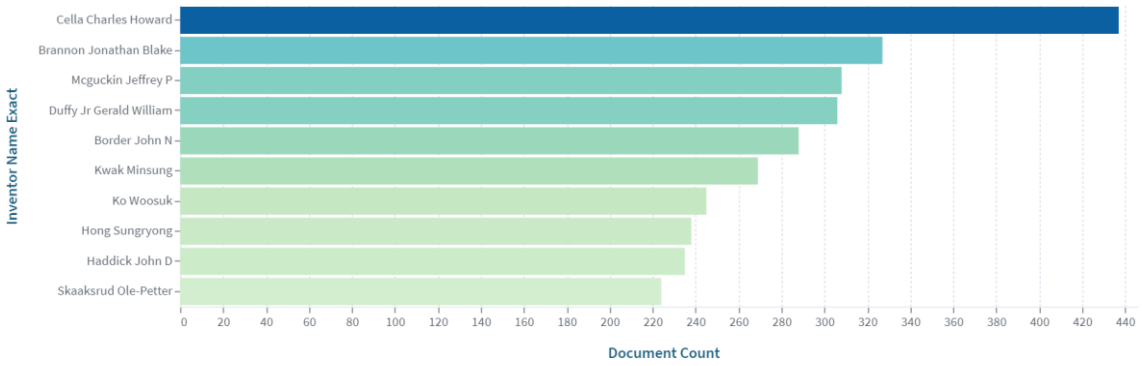
В результате найдены 18073 записи, после фильтрации по годам выпуска 2010–2021 получен набор из 9069 научных публикаций.

2. Результат анализа патентов и научных публикаций.

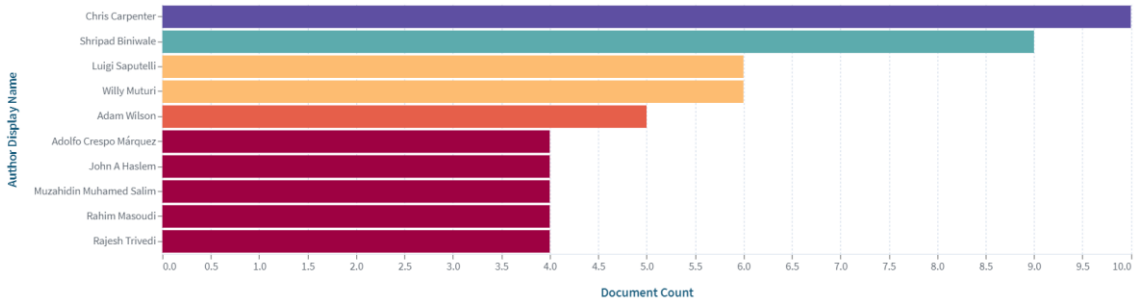
2.1. Самые активные изобретатели и ученые в области APM. Выявление персоналий, вносящих вклад в APM, является важным первым шагом в понимании основных акторов в анализе и установлении значимых партнерских отношений.

Эти графики (рис. 6, 7) содержат данные о наиболее известных и влиятельных ученых и изобретателях, вносящих свой вклад в APM. Выявление этих ключевых лиц помогает определить, кто стоит за исследованиями и изобретениями по APM и способствует развитию партнерских отношений. Исследование участников инновационной системы необходимо, чтобы лучше понять возможности каждого из них, найти и поделиться возможностями, наладить партнерские отношения, минимизировать риск собственных исследований и выявить перспективные траектории сотрудничества.

Рисунки 6 и 7 иллюстрируют данные о ведущих ученых и изобретателях в Lens, которые внесли свой вклад в развитие APM, на основе анализа количества их научных работ, количества патентов, научного или патентного цитирования.

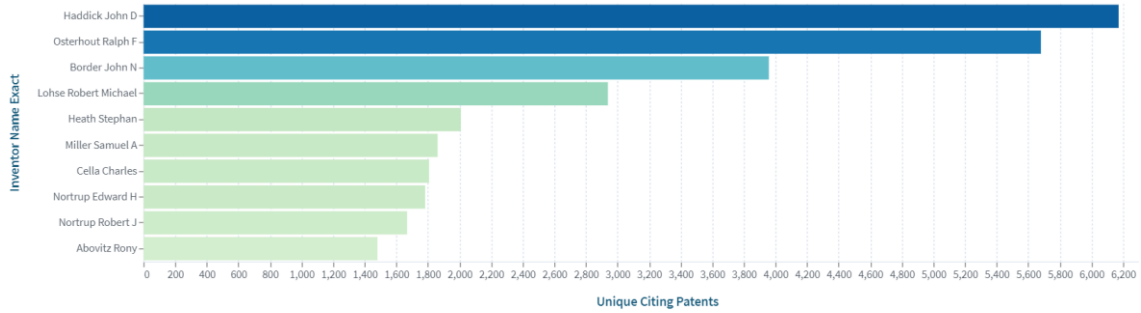


а)

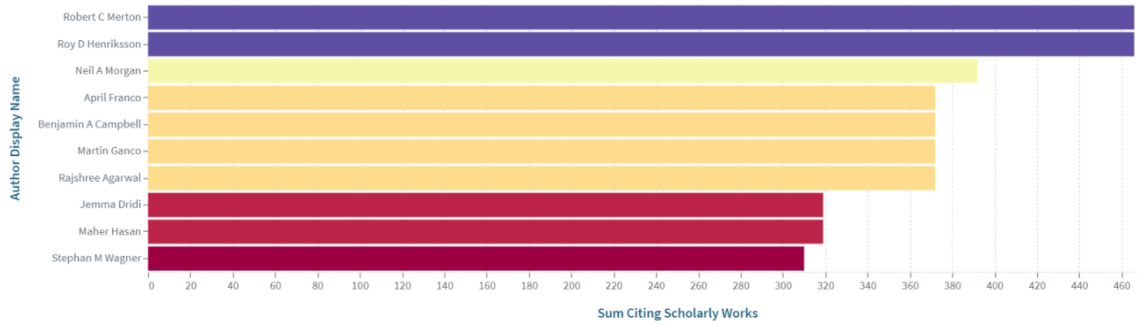


б)

Рисунок 6 – Самые активные (а) изобретатели и (б) ученые в области АРМ



а)



б)

Рисунок 7 – Исследователи с наибольшим числом (а) цитируемых патентов и (б) цитируемых научных работ

2.2. Самые влиятельные организации. Определение того, какие организации являются наиболее влиятельными (рис. 8) в области АРМ, позволяет различным участникам инновационной системы находить друг друга, чтобы лучше понимать возможности каждого и создавать партнерства, минимизировать риски внедрения высокотехнологичных решений и принимать решения по прямому взаимодействию.

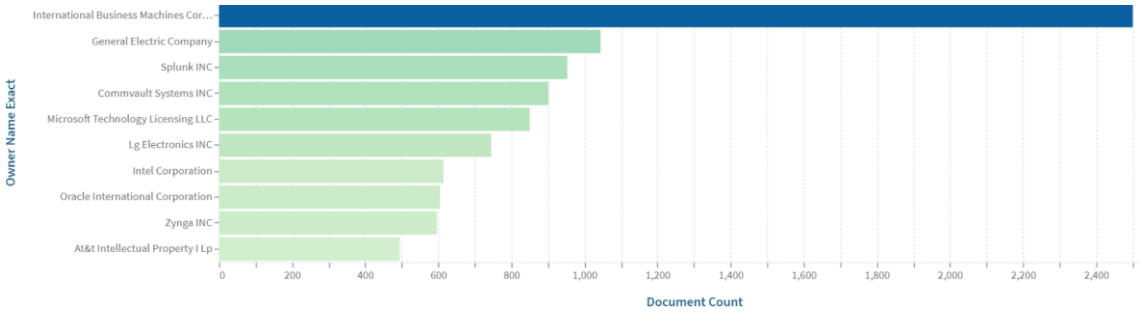


Рисунок 8 – Ведущие организации по количеству патентных документов

2.3. Самые востребованные тематики и области исследования. The Lens представляет собой агрегатор метаданных, объединяющий три уникальных набора контента:

- Microsoft academic;
- CrossRef;
- PubMed

и один базовый инструмент управления. Этот инструмент поддерживает основные функции Lens, которые заключаются в обнаружении, анализе и управлении следующими объектами:

- научные публикации;
- патенты;
- PatSeq – средство для поиска и анализа биологических последовательностей, раскрытых в патентной литературе;
- коллекции – Collections – инструмент управления для динамического или статического отслеживания, мониторинга и анализа коллекции научных работ или патентов [10].

Среди них научные работы обнаруживаются и анализируются с помощью инструментов, обеспечивающих доступ к глобальному корпусу метаданных научной литературы с индексацией цитирования, в основном от Microsoft Academic. Эта платформа использует машинное обучение, семантический вывод и извлечение знаний, чтобы обеспечить эффективное исследование научной информации [11, 12]. В частности, Microsoft Academic использовал содержимое Википедии и анализ графических ссылок, чтобы расширить охват областей исследования (Fields of study – FoS). Начиная с нескольких тысяч высококачественных начальных значений FoS и повторяя несколько циклов анализа связей графа и фильтрации сущностей, можно выявить больше FoS. Затем были просканированы сотни миллионов метаданных академических публикаций, таких как название, ключевые слова, аннотации, чтобы подтвердить существование нового FoS. Основываясь на тематиках (Subjects) и областях исследования FoS, определенных Microsoft Academic в каждой научной работе, можно обобщить наиболее часто встречающиеся FoS (рис. 9) и предметы (рис. 10) в исследованиях АРМ.

Самые частые области исследования состоят из следующих: Business, Finance, Economics, Computer science, Accounting, Return on assets и т. д.

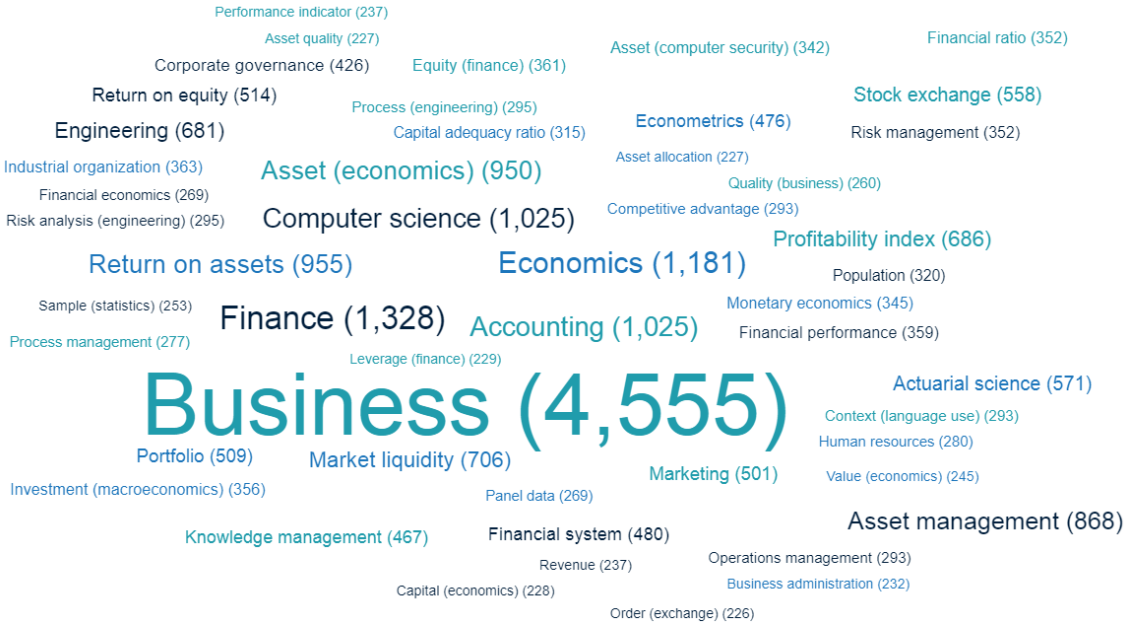


Рисунок 9 – Облако ТОП-50 FoS

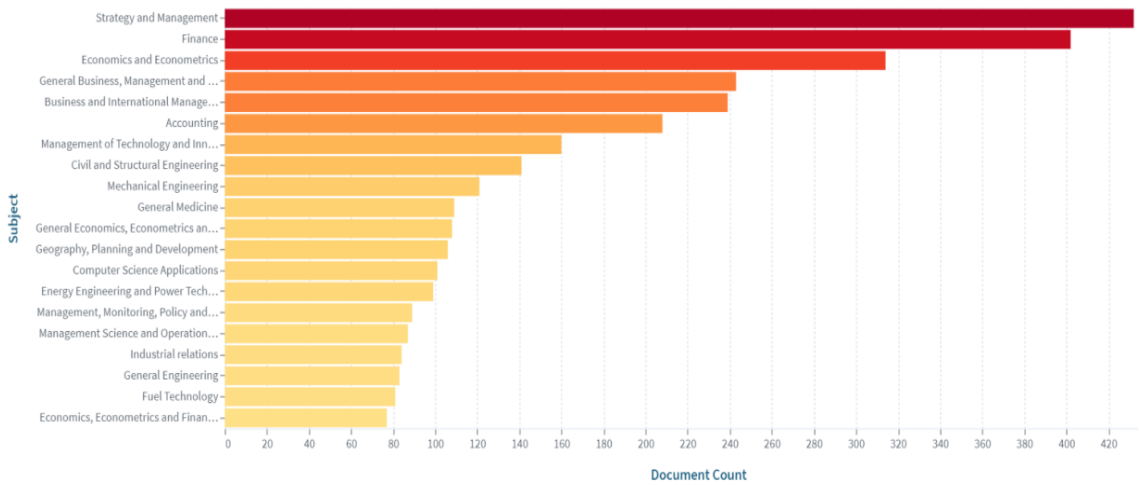


Рисунок 10 – Наиболее частые тематики исследования в научных публикациях по APM

Из рисунка 10 видны наиболее частые тематики исследования в научных публикациях по APM: Strategy and Management; Finance; Economics and Econometrics; General Business, Management and Accounting и т. д.

С другой стороны, ниже перечислены описания 5 самых популярных классов CPC (рис. 11) в патентах по порядку убывания частоты появления:

- G06N20/00: машинное обучение;
- H04L67/10: зависящие от сети механизмы или протоколы связи, поддерживающие сетевые приложения, в которых приложение распределяется по узлам в сети (механизмы мультипрограммирования);
- H04L2209/38: дополнительная информация или приложения, относящиеся к криптографическим механизмам или криптографическим схемам для секретной или безопасной связи (цепочка хеширования или цепочка сертификатов);
- H04L67/12: зависящие от сети механизмы или протоколы связи, поддерживающие сетевые приложения, адаптированы для проприетарных или специализированных сетевых сред, например медицинские сети, сенсорные сети, сети в автомобиле или сети удаленных измерений (сети домашней автоматизации; тотальный заводской контроль, характеризующийся сетевой связью; игры, включающие системы передачи);
- H04L9/3239: криптографические механизмы для секретной или безопасной связи (сетевая архитектура или сетевые протоколы связи для сетевой безопасности или для безопасности беспроводной сети; меры безопасности для защиты компьютеров или компьютерных систем от несанкционированной активности) с участием неключевых хэш-функций, например коды обнаружения модификации (MDC), MD5, SHA или RIPEMD.

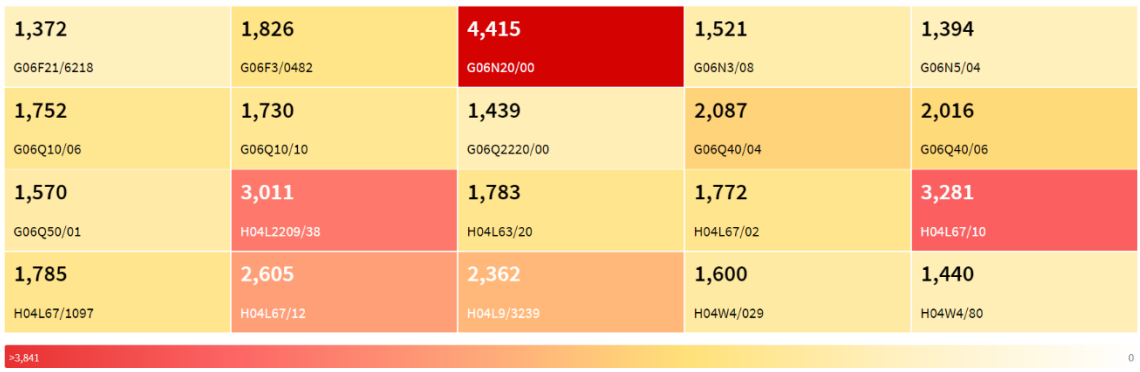


Рисунок 11 – Самые популярные классы CPC

2.4. Выделение тенденции исследования на основе представления на временной шкале. Знание хронологии открытий, изобретений и событий в анализе предметной области АРМ помогает нам понять сигналы во временных тенденциях и интерпретировать стадии развития и уровни зрелости технологий.

Сопоставляя временные рамки открытий (представленных временными рядами научных работ, рис. 12) и изобретений (представленных временными рядами патентов, рис. 13) с политическими, правовыми и нормативными событиями, можно идентифицировать ключевые события, такие как крупные открытия, развитие вспомогательных технологий или основные политические решения.

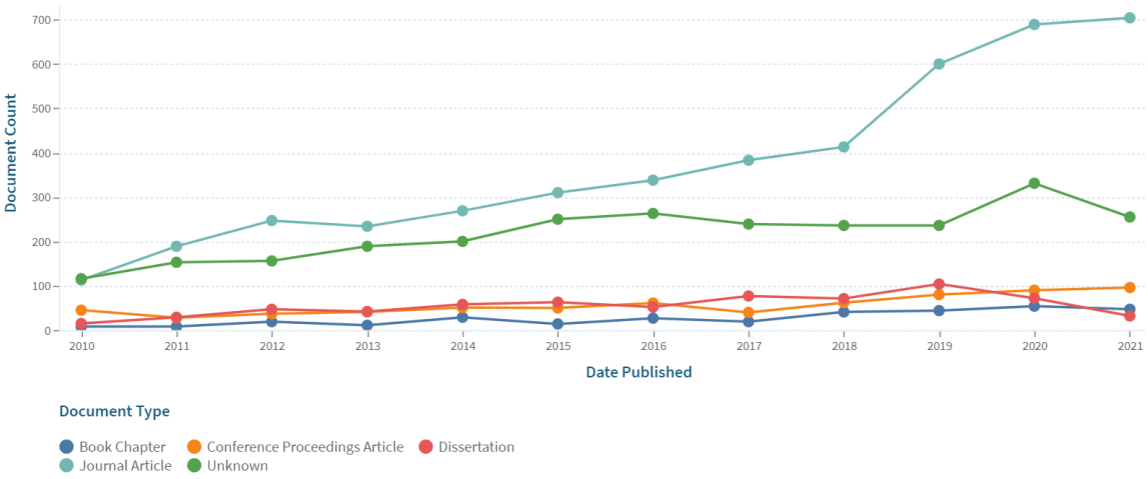


Рисунок 12 – Динамика научных публикаций по годам

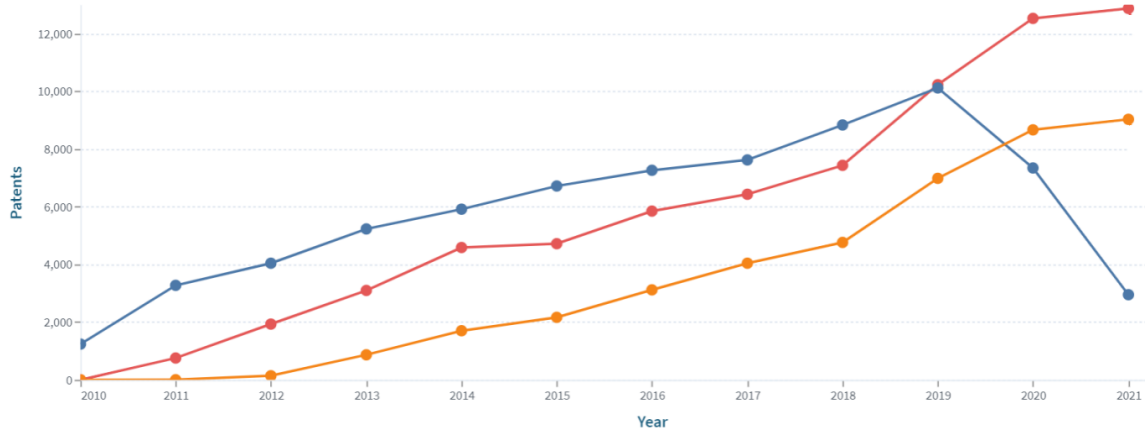


Рисунок 13 – Динамика количества поданных, выданных, опубликованных патентов

Из рисунка 12 видно, что количество публикаций по тематике АРМ возрастает каждый год. Возможно, появились новые области знаний, которые связаны с данной темой и находятся в тренде. Для выявления трендов тематик исследования рассмотрим динамику их появления по годам в научных публикациях [13, 14]. Затем приспособим выраженную динамику моделью линейной регрессии [15]. В частности, коэффициент регрессии (наклон) будет рассчитываться с помощью метода наименьших квадратов (Ordinary least square – OLS). Зависимой переменной оценки является показатель частоты появления тематик, а независимой переменной является год в анализируемом периоде (2010–2021). При этом коэффициент регрессии показывает, насколько в среднем частота анализируемой тематики менялась с каждым наблюдаемым годом. Поэтому для выбора развивающейся тематики исследования следует отобрать положительный коэффициент регрессии ($a \geq 1$) и p-value < 5 %.

Далее для проверки и прогнозирования тенденции изменения частоты тематик необходимо применить метод скользящих средних (была использована функция seasonal_decompose в библиотеке statsmodels [16]). В результате выбраны 9 тематик исследования, которые имеют большой наклон регрессии и сохраняют восходящие тенденции. Все данные динамики частоты, аппроксимирующей линии и линии тенденции выбранных тематик показаны на рисунке 14.

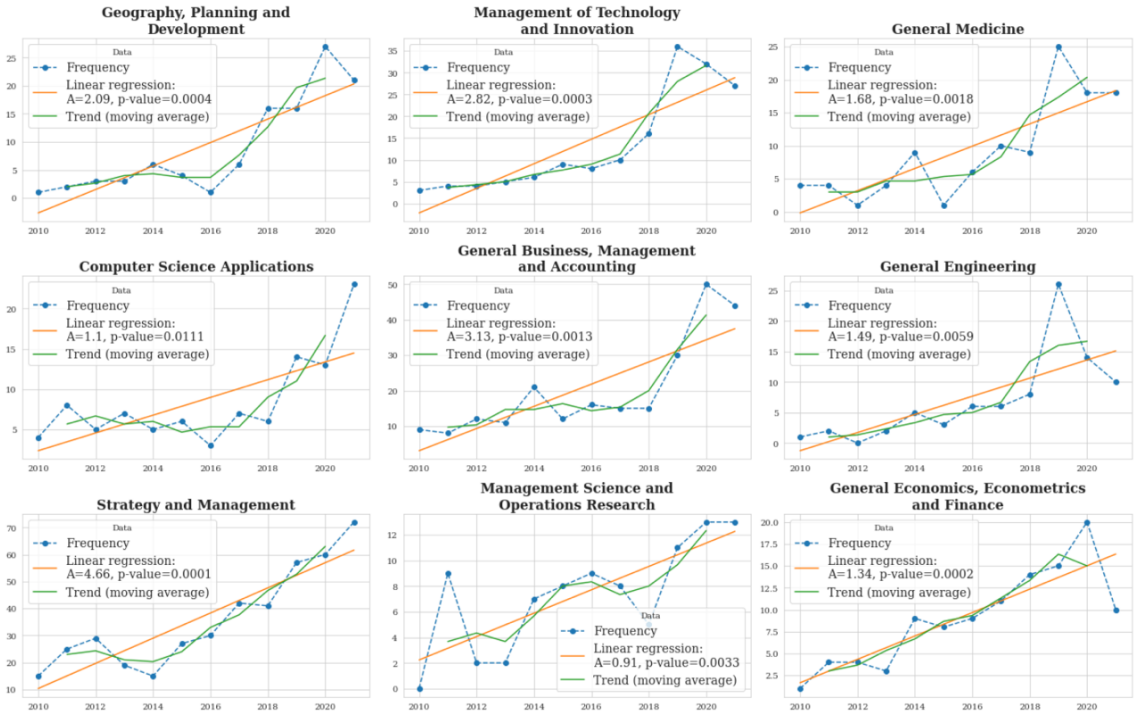


Рисунок 14 – Выявление тематик исследования по тренду появления

С другой стороны, приведена динамика количества поданных, выданных, опубликованных патентов, которая иллюстрирована на рисунке 13. Из рисунка видно, что хотя количество поданных, выданных патентов увеличивается ежегодно, количество патентных заявок падает с 2019 г. Отсюда логично предположить, что инновации, связанные с технологией АРМ, замедляются на последние годы. Однако количество научных публикаций еще сохраняет свой восходящий тренд, поэтому можно сделать вывод о продолжительном активном развитии фундаментальных исследований области АРМ.

2.5. Анализ научных публикаций, цитируемых в патентах по АРМ. В последние 20 лет количество патентных ссылок на научные статьи быстро увеличивается. Это свидетельствует о все более сильном перетекании знаний из академической науки в промышленные инновации. Многие исследования доказали, что склонность патентов к цитированию академических публикаций резко возросла, даже с учетом изменений в объеме и распределении патентов в разных областях [17].

Примечательно, в исследовании [18] авторы исследуют, в какой степени патентные ссылки на статьи могут служить ранними признаками для прогнозирования замедленно распознанных (delayed recognized – DR) знаний в науке, используя сравнительное исследование с контрольной группой, т. е. статьи с мгновенным распознаванием (instant recognition papers). Выходило, что работы с замедленным распознаванием оказывают более сильное и длительное техническое воздействие, чем работы с мгновенным распознаванием.

Более того, важный тип основанных на цитировании ранних признаков пробуждения (awakening) DR-публикаций или признания преждевременных (premature) открытий и преобразующих исследований (transformative research) сосредоточен на патентных ссылках на DR-статьи и научные непатентные источники. В заключение авторы отметили, что половина DR-публикаций по физике, химии, технике и информатике являются прикладными исследованиями или публикуются в прикладных журналах и значительно больше цитируются в патентах, чем другие общенаучные работы [19].

При этом связи между наукой и технологиями широко изучались с использованием ссылок на непатентную литературу или сопоставления авторов и изобретателей. В статье [20] используется латентное распределение Дирихле (LDA) для создания тематических связей между публикациями и патентами на основе семантического содержания документов. Этот подход позволяет обнаруживать тематические совпадения между патентными и научными публикациями, выделяя тематические области, примененные для исследований и технологий.

Таким образом, очень важно анализировать научные работы, на которые ссылаются патенты АРМ. На момент написания данной статьи существует 29943 научных публикации, цитируемых в найденных патентах по АРМ. На рисунке 15 показаны самые влиятельные организации в цитируемых научных публикациях, такие как Stanford University, Massachusetts Institute of Technology (MIT), International Business Machines (IBM), Harvard University, University of Washington (UW) и т. д.

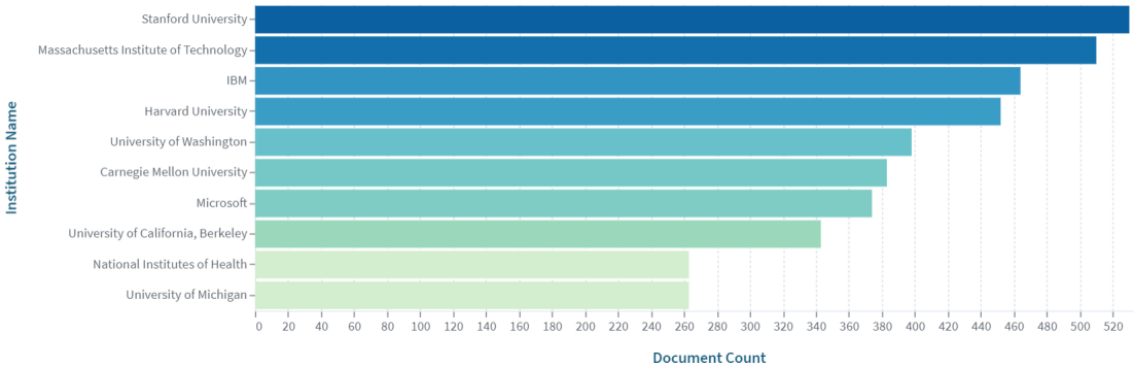


Рисунок 15 – Самые влиятельные организации в цитируемых научных публикациях

На рисунке 16 приведено облако топ-50 областей исследования в цитируемых научных публикациях. Наиболее существенные области исследования включают: Computer Science (computer network, computer security, computer vision), Biology, Medicine, Artificial Intelligence, Engineering, Chemistry и т. д.

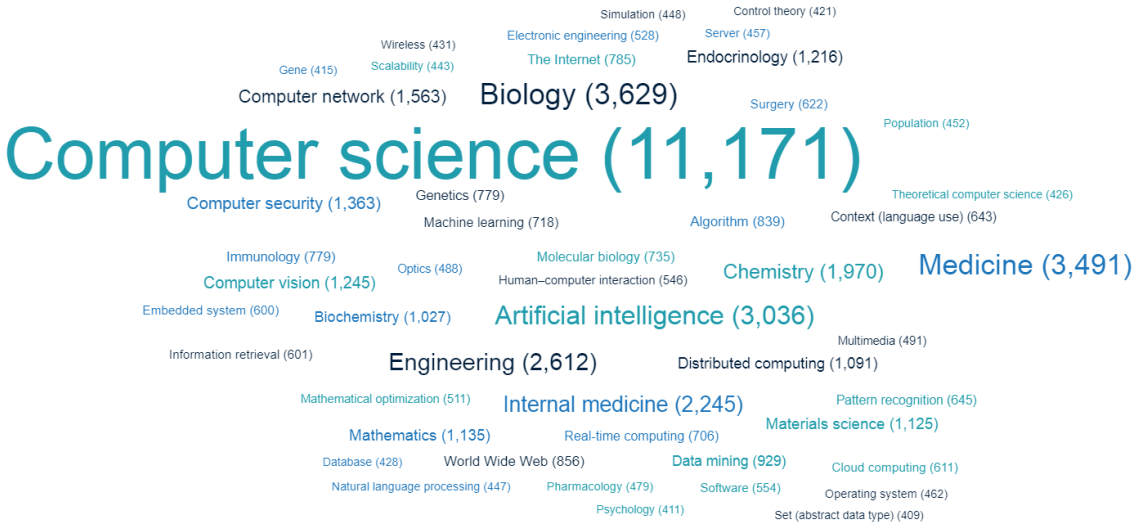


Рисунок 16 – Облако ТОП-50 областей исследования в цитируемых научных публикациях

Причем ТОП-20 тематик исследования в цитируемых научных публикациях показаны на рисунке 17, которые представляют собой: Electrical and Electronic Engineering, General Medicine, Software, General Computer Science (and Applications), Biochemistry, Molecular Biology, Computer Networks and Communications, Theoretical Computer Science и т. д.

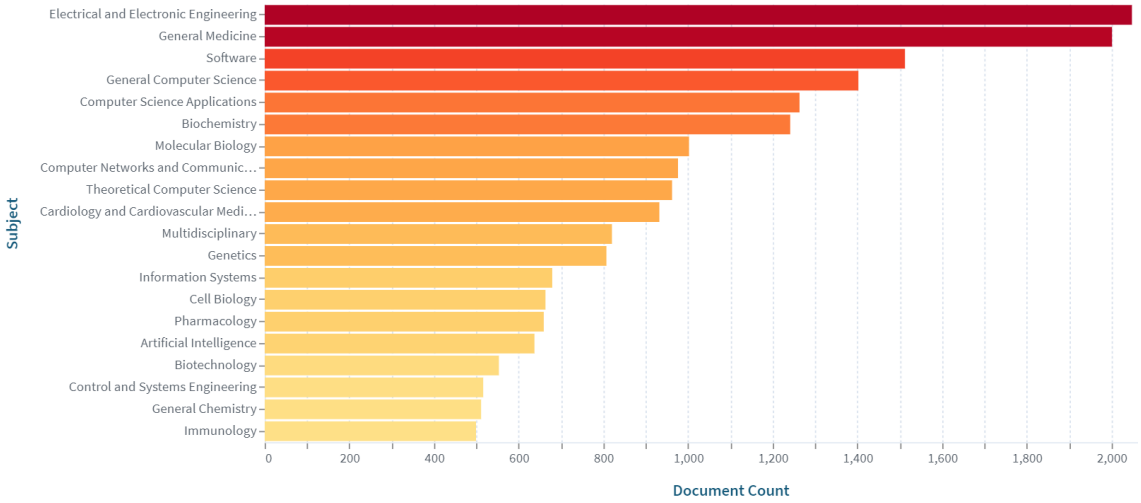


Рисунок 17 – Топ-20 тематик исследования в цитируемых научных публикациях

Заключение. Технологические разработки оказывают существенное влияние на принятие стратегических решений. Раннее осознание возможных грядущих или появляющихся технологических тенденций может привести к усилению конкурентоспособности и рыночного позиционирования предприятий [21, 22]. Однако, если инновационные компании игнорируют появляющиеся технологические разработки, они не смогут полностью реализовать потенциал своих собственных продуктов или технологий.

Подход, ориентированный на данные, повышает эффективность актива, что является дополнительным экономическим преимуществом для эксплуатирующего предприятия. При этом управление эффективностью активов стало основным инструментом цифровой трансформации промышленных компаний. В работе предложен и реализован метод анализа тенденции развития технологий в области АРМ, в результате выделены наиболее влиятельные факторы (изобретатели, ученые, организации), значительные направления исследования и основные тенденции в данной предметной области. Следовательно, необходимо отметить, что основными трендами области АРМ за последние 12 лет (2010–2021) являются следующие:

1. В настоящее время данная область исследования АРМ еще находится в стадии продолжительной разработки научных исследований и довольно активно развивается. При этом обнаружены наиболее частые предметы (Subject) и области исследования (Field of study) в научных публикациях по АРМ. Также выявлены 9 тематик восходящего тренда: Geography, Planning and Development; Management of Technology and Innovation; General Medicine; Computer Science Applications; General Business, Management and Accounting; General Engineering; Strategy and Management; Management Science and Operations Research; General Economics, Econometrics and Finance.

2. Ведущие предприятия по количеству патентов в области включают: IBM Corporation, General Electric Company, Splunk INC, Commvault Systems INC, Microsoft Technology Licensing LLC и др. Причем наиболее значительные классы патентования за 12 лет: машинное обучение; зависящие от сети механизмы или протоколы связи, поддерживающие сетевые приложения; приложения, относящиеся к криптографическим механизмам или криптографическим схемам для секретной или безопасной связи и т. д. Однако за последние 3 года (2019–2021) эти классы патентования показывают убывающую тенденцию в патентных заявках.

Кроме того, при анализе научных публикаций, цитируемых в АРМ-патентах, также выявлены наиболее влияющие на инновационное развитие АРМ-организации, области исследования и тематики исследования.

Библиографический список

1. Управление эффективностью активов: новый взгляд на преимущества Asset Performance Management 4.0. – 2020. – Режим доступа: https://sapr.ru/article/26153, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 01.02.2022).

2. Ferket, J. Asset performance management 4.0 internet of things IoT enabled condition monitoring, a story from a digital maintenance service provider / J. Ferket // Society of Petroleum Engineers – Abu Dhabi International Petroleum Exhibition and Conference 2018, ADIPEC 2018. – 2019.

3. Kizim, A. V. On Systemological Approach to Intelligent Decision-Making Support in Industrial Cyber-Physical Systems / A. V. Kizim, A. G. Kravets // Studies in Systems, Decision and Control. – 2020. – P. 167–183.

4. Karar, A.N., Labib A., Jones D.F. A conceptual framework for an agile asset performance management process / A. N. Karar, A. Labib, D. F. Jones // Journal of Quality in Maintenance Engineering. – 2021.

5. Тихомиров, Л. И. Информационные центры управления производственными активами - ключевой элемент повышения эффективности ТОиР / Л. И. Тихомиров, В. В. Лехтцинд // Автоматизация в промышленности. – 2021. – Т. 12. – С. 28–31.

6. Князева, Н. В. Разработка единой информационной модели промышленного предприятия для повышения эффективности управления его непроизводственными активами / Н. В. Князева, В. А. Епифанов // Управленческий учет. – 2021. – Т. 12, № 2. – С. 418–426.

7. Yusupbekov, N. Concepts and Methods of “Digital Twins” Models Creation in Industrial Asset Performance Management Systems / N. Yusupbekov et al. // Advances in Intelligent Systems and Computing. – 2021. – P. 1589–1595.

8. Shcherbakov, M. A method and IR4I index indicating the readiness of business processes for data science solutions / M. Shcherbakov, P. P. Groumpos, A. Kravets // Communications in Computer and Information Science. – 2017. – P. 21–34.

9. Wan, S. Asset Performance Management for Power Grids / S. Wan // Energy Procedia. – 2017. – P. 611–616.

10. About The Lens. – Режим доступа: https://about.lens.org/, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 30.01.2022).

11. Sinha, A. и др. An Overview of Microsoft Academic Service (MAS) and Applications / A. Sinha et al. // Proceedings of the 24th International Conference on World Wide Web. – New York, NY, USA: ACM, 2015. – P. 243–246.

12. Wang, K. A Review of Microsoft Academic Services for Science of Science Studies / K. Wang et al. // Frontiers in Big Data. – 2019. – Vol. 2.

13. Нгуен, Т. В. Анализ и прогноз тенденций использования терминов в компьютерных науках на основе нейросетевых моделей / Т. В. Нгуен, А. Г. Кравец, К. Х. Т. Зыонг // Вестник компьютерных и информационных технологий. – 2021. – Т. 18, № 2. – С. 24–38.

14. Нгуен, Т. В. Оценка и прогнозирование тенденций развития научных исследований на основе библиометрического анализа публикаций / Т. В. Нгуен, А. Г. Кравец // Информационные технологии. – 2021. – Т. 27, № 4. – С. 195–201.

15. Ordinary least squares Linear Regression. – Режим доступа: https://scikit-learn.org/stable/modules /generated/sklearn.linear_model.LinearRegression.html, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 30.01.2022).

16. Statsmodels.tsa.seasonal.seasonal_decompose. – Режим доступа: https://www.statsmodels.org/stable /generated/statsmodels.tsa.seasonal.seasonal_decompose.html, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 30.01.2022).

17. Wei, F. Decreasing the noise of scientific citations in patents to measure knowledge flow / F. Wei et al. // 17th International Conference on Scientometrics and Informetrics, ISSI 2019 – Proceedings. – 2019. – P. 1662–1669.

18. Du, J. Paper-patent citation linkages as early signs for predicting delayed recognized knowledge: Macro and micro evidence / J. Du et al. // Journal of Informetrics. – 2020. – Vol. 14, № 2.

19. van Raan, A. F. J. Sleeping beauties cited in patents: Is there also a dormitory of inventions? / A. F. J. van Raan // Scientometrics. – 2017. – Vol. 110, № 3. – P. 1123–1156.

20. Suominen, A. Exploration of Science and Technology Interaction: A Case Study on Taxol / A. Suominen, S. Ranaei, O. Dedehayir // IEEE Transactions on Engineering Management. – 2021. – Vol. 68, no. 6. – P. 1786–1801.

21. Кравец, А. Г. Формальные метрики для автоматизированной оценки изобретений / А. Г. Кравец, М. С. Легенченко // Прикаспийский журнал: управление и высокие технологии. – 2017. – № 3 (39). – С. 8–19.

22. Кравец, А. Г. Предсказательное моделирование трендов технологического развития / А. Г. Кравец, Н. А. Сальникова // Известия Санкт-Петербургского государственного технологического института (технического университета). – 2020. – Т. 55, № 81. – С. 103–108.

References

1. *Upravlenie effektivnostyu aktivov: novyy vzglyad na preimushchestva Asset Performance Management 4.0* [Asset Performance Management: A New Look at the Benefits of Asset Performance Management 4.0]. Available at: https://sapr.ru/article/26153 (accessed 01.02.2022).

2. Ferket, J. Asset performance management 4.0 internet of things IoT enabled condition monitoring, a story from a digital maintenance service provider. *Society of Petroleum Engineers – Abu Dhabi International Petroleum Exhibition and Conference 2018, ADIPEC 2018*. 2019.

3. Kizim, A. V., Kravets, A. G. On Systemological Approach to Intelligent Decision-Making Support in Industrial Cyber-Physical Systems. *Studies in Systems, Decision and Control*, 2020, pp. 167–183.

4. Karar, A. N., Labib, A., Jones, D. F. A conceptual framework for an agile asset performance management process. *Journal of Quality in Maintenance Engineering*, 2021.

5. Tihomirov, L. I., Lehtcind, V. V. Informatsionnye tsentry upravleniya proizvodstvennymi aktivami – klyuchevoy element povysheniya effektivnosti TOiR [Information centers for managing production assets are a key element in increasing the efficiency of maintenance and repair]. *Avtomatizatsiya v promyshlennosti* [Automation in Industry], 2021, vol. 12, pp. 28–31.

6. Knyazeva, N. V., Epifanov, V. A. Razrabotka edinoy informatsionnoy modeli promyshlennogo predpriyatiya dlya povysheniya effektivnosti upravleniya ego neproizvodstvennymi aktivami [Development of a unified information model of an industrial enterprise to improve the efficiency of managing its non-productive assets]. *Upravlencheskiy uchet* [Management Accounting], 2021, vol. 12, no. 2, pp. 418–426.

7. Yusupbekov, N. et al. Concepts and Methods of “Digital Twins” Models Creation in Industrial Asset Performance Management Systems. *Advances in Intelligent Systems and Computing*, 2021, pp. 1589–1595.

8. Shcherbakov, M., Groumpos, P.P., Kravets, A. A method and IR4I index indicating the readiness of business

processes for data science solutions. *Communications in Computer and Information Science*, 2017, pp. 21–34.

9. Wan, S. Asset Performance Management for Power Grids. *Energy Procedia*, 2017, pp. 611–616.

10. *About The Lens*. Available at: https://about.lens.org/ (accessed 30.01.2022).

11. Sinha, A. et al. An Overview of Microsoft Academic Service (MAS) and Applications. *Proceedings of the 24th International Conference on World Wide Web*. New York, NY, USA, ACM, 2015, pp. 243–246.

12. Wang, K. et al. A Review of Microsoft Academic Services for Science of Science Studies. *Frontiers in Big Data*, 2019, vol. 2.

13. Nguyen, T. V., Kravets, A. G., Duong, Q. H. T. Analiz i prognoz tendentsiy ispolzovaniya terminov v kompyuternykh naukakh na osnove neyrosetevykh modeley [Analysis and trend prediction of using terms in computer science based on neural network models]. *Vestnik kompyuternykh i informatsionnykh tehnologiy* [Bulletin of Computer and Information Technologies], 2021, vol. 18, no. 2, pp. 24–38.

14. Nguyen, T. V., Kravets, A. G. Otsenka i prognozirovanie tendentsiy razvitiya nauchnykh issledovaniy na osnove bibliometricheskogo analiza publikatsiy [Assessment and forecasting of scientific research development trends based on bibliometric analysis of publications]. *Informatsionnye tekhnologii* [Information Technology], 2021, vol. 27, no. 4, pp. 195–201.

15. *Ordinary least squares Linear Regression*. Available at: https://scikit-learn.org/stable/modules/generated /sklearn.linear_model.LinearRegression.html (accessed 30.01.2022).

16. *Statsmodels.tsa.seasonal.seasonal_decompose*. Available at: https://www.statsmodels.org/stable/generated /statsmodels.tsa.seasonal.seasonal_decompose.html (accessed 30.01.2022).

17. Wei, F. et al. Decreasing the noise of scientific citations in patents to measure knowledge flow. *17th International Conference on Scientometrics and Informetrics, ISSI 2019 – Proceedings*, 2019, pp. 1662–1669.

18. Du, J. et al. Paper-patent citation linkages as early signs for predicting delayed recognized knowledge: Macro and micro evidence. *Journal of Informetrics*, 2020, vol. 14, no. 2.

19. van Raan, A. F. J. Sleeping beauties cited in patents: Is there also a dormitory of inventions? *Scientometrics*, 2017, vol. 110, no. 3, pp. 1123–1156.

20. Suominen, A., Ranaei, S., Dedehayir, O. Exploration of Science and Technology Interaction: A Case Study on Taxol. *IEEE Transactions on Engineering Management*, 2021, vol. 68, no. 6, pp. 1786–1801.

21. Kravets, A.,G., Legenchenko, M.,S. Formalnye metriki dlya avtomatizirovannoyy otsenki izobreteniy [Formal Metrics for Automated Evaluation of Inventions]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2017, no. 3 (39), pp. 8–19.

22. Kravets, A. G., Salnikova, N. A. Predskazatelnoe modelirovanie trendov tekhnologicheskogo razvitiya [Predictive modeling of technological development trends]. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo tekhnologicheskogo instituta (tekhnicheskogo universiteta)* [Proceedings of the St. Petersburg State Technological Institute (Technical University)], 2020, vol. 55, no. 81, pp. 103–108.

DOI 10.54398/2074-1707_2022_1_53
УДК 004.001

СИСТЕМНОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ ЗАТРАТ ПРОЦЕССА УПРАВЛЕНИЯ ОПЕРАТИВНО-ВЫЕЗДНЫМИ БРИГАДАМИ ЭЛЕКТРОСЕТЕВОЙ КОМПАНИИ

Статья поступила в редакцию 06.01.2022, в окончательном варианте – 29.01.2022.

Кинжалиева Алия Рахметовна, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 16,
аспирант, ORCID: 0000-0002-6243-5750, e-mail: satobalova@mail.ru

Ханова Анна Алексеевна, Астраханский государственный технический университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 16,
доктор технических наук, доцент, ORCID: 0000-0003-2693-8876, e-mail: akhanova@mail.ru

Реализация задачи обеспечения надежности оказания услуг электросетевыми компаниями при минимальном уровне затрат во многом зависит от сокращения продолжительности перерыва электроснабжения потребителей, что может быть достигнуто путем совершенствования процесса управления дежурным персоналом оперативно-выездных бригад. В связи с этим задача повышения эффективности процесса управления оперативно-выездными бригадами при возникновении аварий и технологических нарушений в электросетевых компаниях является актуальной. Построены временные графики, рассмотрена структура и динамика затрат процесса управления оперативно-выездными бригадами. Представлено формальное описание процесса формирования затрат на эксплуатацию собственных и привлеченных ресурсов при возникновении аварийных ситуаций в электросетевом комплексе. Описаны уравнения, определяющие темпы потоков и уровни затрат, формируемых при эксплуатации ресурсов. Разработана и предложена прикладная базовая модель системной динамики на основе диаграммы потоков Форрестера для создания имитационных моделей распределительных электросетевых компаний. Представлена имитационная модель системной динамики, разработанная в программном продукте Anylogic. Разработана схема проведения эксперимента с имитационной моделью в нотации EPC. Детально описаны результаты экспериментов с моделью, выявлены оптимальные характеристики в разрезе численности персонала и общего количества задействованной техники.

Ключевые слова: системная динамика, имитационное моделирование, затраты, электросетевая компания, оперативно-выездные бригады, авария, управление, персонал, ресурсы

SYSTEM MODELING OF THE COST DYNAMICS OF THE MANAGEMENT PROCESS OF OPERATIONAL-FIELD TEAMS OF AN ELECTRIC GRID COMPANY

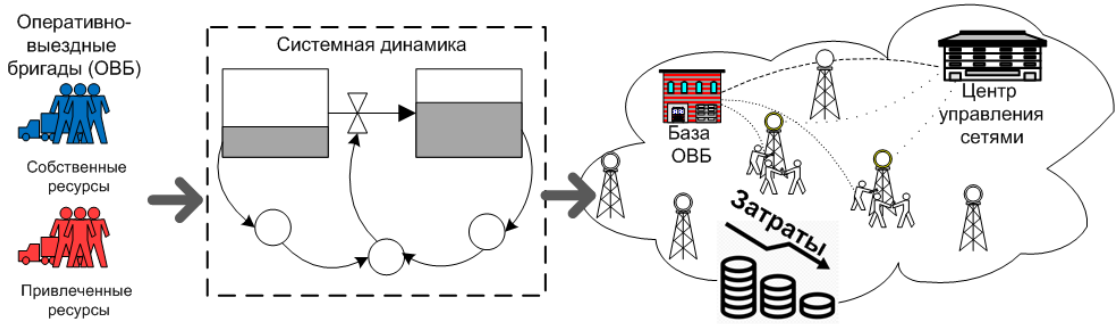
The article was received by the editorial board on 06.01.2022, in the final version – 29.01.2022.

Kinzhalieva Aliya R., Astrakhan State Technical University, 16 Tatishchev St., Astrakhan, 414056, Russia Federation,
post-graduate student, ORCID: 0000-0002-6243-5750, e-mail: satobalova@mail.ru
Khanova Anna A., Astrakhan State Technical University, 16 Tatishchev St., Astrakhan, 414056, Russia Federation,
Doct. Sci. (Engineering), Associate Professor, ORCID: 0000-0003-2693-8876, e-mail: akhanova@mail.ru

The implementation of the task of ensuring the reliability of the provision of services by electric grid companies at a minimum level of costs largely depends on reducing the duration of the interruption of power supply to consumers, which can be achieved by improving the management process of the on-duty personnel of operational field teams. In this regard, the task of improving the efficiency of the management process of operational field teams in the event of accidents and technological violations in electric grid companies is urgent. Time schedules are constructed, the structure and cost dynamics of the management process of operational and field teams are considered. A formal description of the process of forming costs for the operation of own and attracted resources in the event of emergency situations in the power grid complex is presented. The equations determining the rates of flows and the levels of costs generated during the exploitation of resources are described. An applied basic model of system dynamics based on the Forrester flow diagram has been developed and proposed to create simulation models of distribution grid companies. A simulation model of system dynamics developed in the Anylogic software product is presented. A scheme for conducting an experiment with a simulation model in EPC notation has been developed. The results of experiments with the model are described in detail, optimal characteristics are identified in terms of the number of personnel and the total number of equipment involved.

Keywords: system dynamics, simulation modeling, costs, electric grid company, field teams, accident, management, personnel, resources

Graphical annotation (Графическая аннотация)



Введение. Современное общество все больше зависит от надежного электроснабжения для обеспечения уровня функциональности и степени удовлетворенности основных потребностей. Как следствие, бесперебойное электроснабжение имеет решающее значение для общества, а электросетевые компании становятся одной из важнейших инфраструктур общества [1], определяемых как физические и логические системы, необходимые для социального обеспечения [2]. Достижение данной задачи оценивается уровнем надежности оказываемых услуг [3]. На указанный показатель непосредственно влияет средняя длительность перерыва электроснабжения потребителей [4]. В связи с этим возникает необходимость поиска путей снижения продолжительности прекращений передачи электроэнергии [5]. Особое внимание следует уделить вопросу организации процесса управления бригадами при возникновении аварий и технологических нарушений, от качества которой зависят сроки выполнения работы по устранению аварий и технологических нарушений, что, в конечном счете, влияет на общую удовлетворенность потребителей [6]. С другой стороны, рациональное использование ресурсов позволяет оптимизировать затраты, связанные с их использованием [7].

Процесс управления оперативно-выездными бригадами при возникновении аварий и технологических нарушений в электросетевых компаниях является сложным, с большим числом возможных ситуаций и состояний. В работах российских и зарубежных авторов для управления и оптимизации процессов выполнения ремонтных работ описано применение методов искусственного интеллекта [8–9], включая нейронные сети [10–11], теорию нечетких множеств [12–13],

и имитационного моделирования [14–16]. При этом стохастический характер возникновения аварийных ситуаций, на который влияют особенности работы оперативно-выездных бригад, распределения материально-технических ресурсов и возникновения затрат подразумевает целесообразность выбора метода имитационного моделирования в качестве инструмента исследования [17]. Среди известных методов построения имитационных моделей (дискретно-событийный, агентный или системно-динамический) необходимо выбрать подход к моделированию, конечным итогом которого должно стать не предсказание или предугадывание будущих экономических ситуаций, но понимание сути динамики исследуемой системы [18–19]. Именно имитационные модели системной динамики позволят обойти ряд ограничений в исследовании процесса управления оперативно-выездными бригадами, вызванных невозможностью на практике получить информацию о данном процессе в различных ситуациях, когда параметры системы и среды меняются во времени и ряд процессов управления можно описать только приближенно [20–21].

Целью работы является повышение эффективности процесса управления оперативно-выездными бригадами (ОВБ) электросетевых компаний посредством системного моделирования динамики затрат.

Материалы и методы решения задачи управления затратами на основе модели системной динамики. Известно, что в изучаемой системе на объектах электросетевой компании в некоторые моменты времени t_n $(n = \overline{1,M})$ по пуассоновскому закону с заданной интенсивностью возникает поток из M аварий h_n длительностью T_n (рис. 1а). В случае получения оповещения о возникновении аварии центром управления сетями (ЦУС) устанавливается категория сложности аварии, а затем определяется наличие (достаточность) на пунктах расположения ОВБ собственных ресурсов (СР), а именно численность дежурного персонала и количество техники. Пусть общая численность дежурного персонала – A чел., общее количество техники, которое может быть задействовано, – B ед. В зависимости от категории сложности аварии и, соответственно, регламентируемого времени на ее устранение центром управления сетями формируется ОВБ из a чел. ($a \in A$) и укомплектовывается необходимой техникой из b ед. ($b \in B$). В случае нехватки собственных ресурсов на место возникновения аварии направляется весь имеющийся резерв ОВБ, а также задействуются привлеченные ресурсы (ПР) – персонал и техника подрядной организации. При этом ОВБ формируется из СР и ПР. В случае когда все собственные ресурсы задействованы на аварийно-восстановительных работах, привлекается минимально необходимый состав бригады подрядной организации.

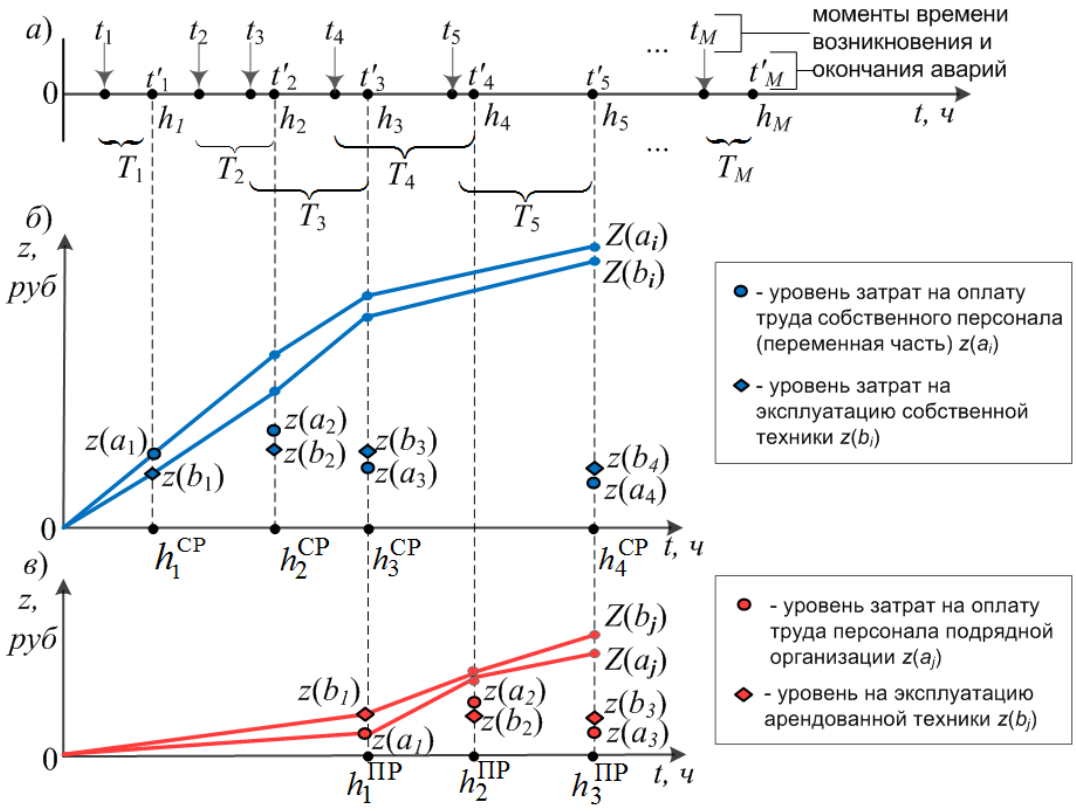


Рисунок 1 – Временные графики процесса управления затратами при возникновении аварий

Для аварий $h^{\text{СР}} \subset h$, на устранение которых задействуются собственные ресурсы, будем использовать счетчик $i = \overline{1,K}, K \le M$, а для аварий $h^{\text{ПР}} \subset h$, на устранение которых задействуются привлеченные ресурсы, – счетчик $j = \overline{1,L}$, $L \le M$. Рассмотрим ситуацию (рис. 2а), когда поочередно возникают аварии h_1 и h_2 $(n = \overline{1,M})$, на устранение которых задействованы только собственные ресурсы. Соответственно, $h_1 = h_1^{\text{СР}}$ и $h_2 = h_2^{\text{СР}}$. Допустим, в тот момент, когда основная часть дежурного персонала и техники еще задействованы при устранении аварий h_1 и h_2, возникает авария h_3. На ее устранение ввиду недостаточности собственных ресурсов требуется дополнительно задействовать привлеченные ресурсы $h_3 = h_3^{\text{СР}} \cup h_1^{\text{ПР}}$. При возникновении аварии h_4 весь имеющийся резерв собственных ресурсов уже задействован на аварийно-восстановительных работах, поэтому для устранения этой аварии задействуются только привлеченные ресурсы $h_4 = h_2^{\text{ПР}}$. Далее часть собственных ресурсов высвобождается, возникает авария $h_5 = h_4^{\text{СР}} \cup h_3^{\text{ПР}}$, для устранения которой задействуются как собственные, так и привлеченные ресурсы. Рассмотрим на данном примере, как отражаются и накапливаются затраты, связанные с использованием ресурсов. На рисунке 2б дано графическое представление динамики постоянной части затрат на оплату труда $z(A)$, которая определяется исходя из дневной тарифной ставки, устанавливаемой дежурному персоналу, независимо от того, привлекался персонал к аварийно-восстановительных работам или нет, а также общей величины постоянной части затрат на оплату труда $Z(A)$.

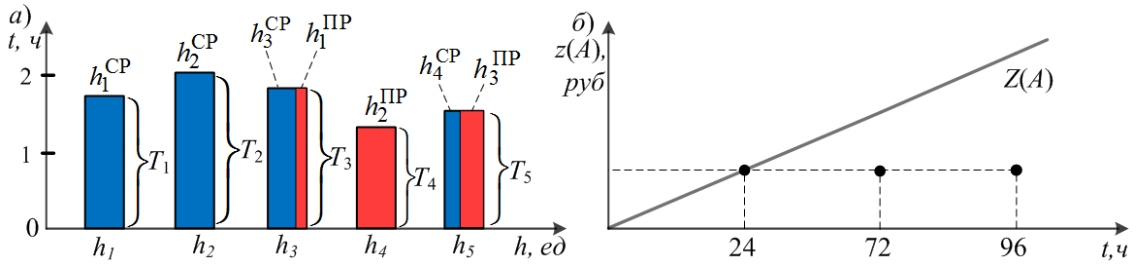


Рисунок 2 – Структура, продолжительность и динамика затрат при возникновении аварий

На рисунке 1б представлено, как в каждый момент времени $t_n' = t_n + T_n$ по окончании устранения аварий за счет собственных ресурсов $h_i^{\text{СР}}$ $(i = 1,\ldots,4)$ зафиксированы трудовые затраты $z(a_i)$ и затраты на использование техники $z(b_i)$, при этом кривые $Z(a_i)$ и $Z(b_i)$ отображают их общую величину. Аналогично на рисунке 1в представлено, как в каждый момент времени t_n' по окончании устранения аварий за счет привлеченных ресурсов $h_j^{\text{ПР}}$ $(j = 1,\ldots,3)$ зафиксированы трудовые затраты $z(a_j)$ и затраты на аренду техники $z(b_j)$, при этом кривые $Z(a_j)$ и $Z(b_j)$ отображают их общую величину. Далее перейдем к концептуальному описанию системы формирования затрат.

Построение системно-динамической модели начнем с создания концептуального описания системы в виде диаграммы потоков Форрестера [22]. Рисунок 3 содержит следующие основные элементы: уровни, изображенные в виде блоков 1–9; потоки, изображенные в виде стрелок и отображающие направления движения от одного уровня к другому; функции решений, изображенные в виде вентилей I–V, используемые для обозначения темпов потоков.

На диаграмме (рис. 3) представлены два основных потока затрат, образующихся за счет эксплуатации СР и ПР. Данные потоки аналогичны друг другу и разбиваются на потоки затрат, связанные с оплатой труда и эксплуатацией техники. По принципу сообщающихся сосудов затраты, накопленные на предыдущих уровнях, влияют на величину затрат на последующих уровнях. Величина затрат на начальном уровне каждого потока зависит от темпа потока затрат (на схеме темпы представлены в виде вентилей I–V). На темпы потоков затрат непосредственно влияют: количество задействованных ресурсов (персонал и техника) (A, a_i, a_j, b_i, b_j), тарифные ставки (S_1, S_2, S_3, S_4, S_5). При этом переменная A представлена в виде константы, поскольку общая численность дежурного персонала является неизменной. В представленной модели отражена упрощенная схема расчета затрат на оплату труда собственного и привлеченного персонала и затрат, связанных с эксплуатацией и арендой техники.

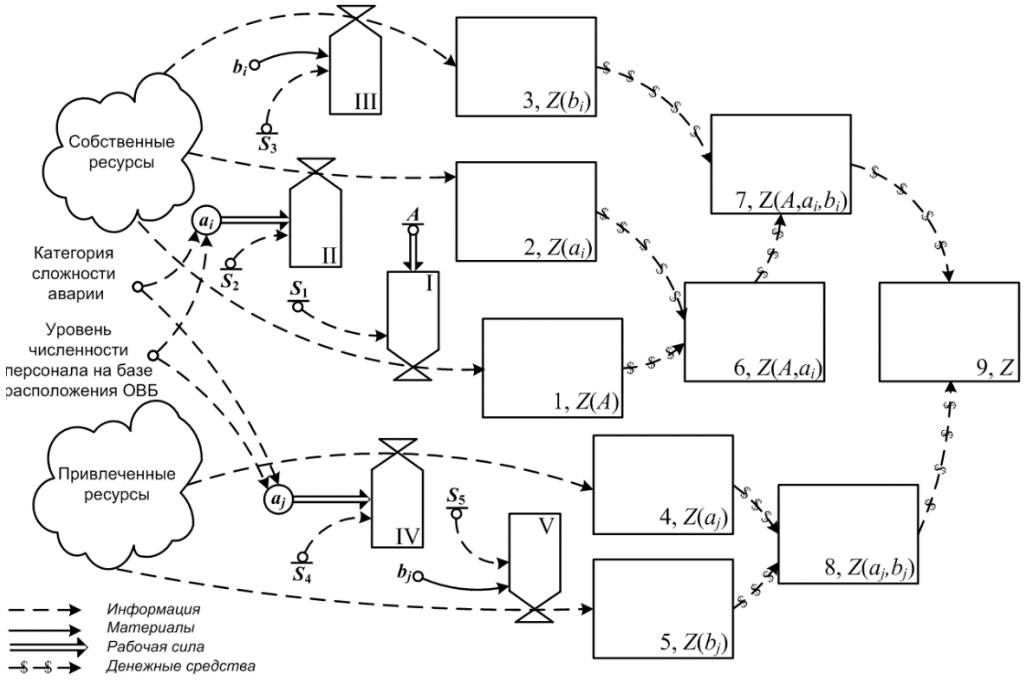


Рисунок 3 – Диаграмма потоков затрат, связанных с устранением аварий

Величина затрат на оплату труда собственного персонала (постоянная часть) z(A) определяется за каждые полные сутки модельного времени m ($m \in Y$, где $Y \leq 365$) и рассчитывается по следующей формуле (рис. 3, вентиль I):

$$z(A) = A \times S_1, \quad (1)$$

где S_1 – размер постоянной части (тарифной ставки) заработной платы собственного персонала, руб.

Общая величина затрат на оплату труда собственного персонала (постоянная часть) Z(A) определяется как сумма затрат на оплату труда собственного персонала (постоянная часть) z(A) и рассчитывается по следующей формуле (рис. 3, уровень 1):

$$Z(A) = \sum_{m=1}^{Y} z(A), \quad (2)$$

Величина затрат на оплату труда собственного персонала (переменной части) z(a_i) рассчитывается по следующей формуле (рис. 3, вентиль II):

$$z(a_i) = T_i \times a_i \times S_2, \quad (3)$$

где S_2 – размер ставки, начисляемой собственному персоналу a_i, задействованного на время T_i для устранения i-й аварии, руб.

Общая величина затрат на оплату труда собственного персонала (переменной части) Z(a_i) рассчитывается по следующей формуле (рис. 3, уровень 2):

$$Z(a_i) = \sum_{i=1}^{k} z(a_i)\,, \quad (4)$$

где k – количество аварий, возникших за m суток модельного времени, устраняемых собственными силами, $i = \overline{1,k}$.

Величина затрат на эксплуатацию собственной техники z(b_i) рассчитывается по следующей формуле (рис. 3, вентиль III):

$$z(b_i) = T_i \times b_i \times S_3, \quad (5)$$

где S_3 – размер ставки за 1 маш.*час (модельного времени) работы собственной техники b_i, задействованной на время T_i для устранения i-й аварии, руб.

Общая величина затрат на эксплуатацию собственной техники Z(b_i) рассчитывается по следующей формуле (рис. 3, уровень 3):

$$Z(b_i) = \sum_{i=1}^{k} z(b_i)\,. \quad (6)$$

Величина затрат на оплату труда персонала подрядной организации $z(a_j)$ рассчитывается по следующей формуле (рис. 3, вентиль IV):

$$z(a_j) = T_j \times a_j \times S_4, \quad (7)$$

где S_4 – размер ставки, начисляемой персоналу подрядной организации a_j, задействованному на время T_j для устранения j-й аварии, руб.

Общая величина затрат на оплату труда персонала подрядной организации $Z(a_j)$ рассчитывается по следующей формуле (рис. 3, уровень 4):

$$Z(a_j) = \sum_{j=1}^{l} z(a_j), \quad (8)$$

где l – количество аварий, возникших за m суток модельного времени, для устранения которых привлекается подрядная организация, $j = \overline{1,l}$.

Величина затрат на эксплуатацию арендованной техники $z(b_j)$ рассчитывается по следующей формуле (рис. 3, вентиль V):

$$z(b_j) = T_j \times b_j \times S_5, \quad (9)$$

где S_5 – размер ставки за 1 маш.*час (модельного времени) работы арендованной техники b_j, задействованной на время T_j для устранения j-й аварии, руб.

Общая величина затрат на эксплуатацию арендованной техники $Z(b_j)$ рассчитывается по следующей формуле (рис. 3, уровень 5):

$$Z(b_j) = \sum_{j=1}^{l} z(b_j) \quad (10)$$

Совокупная величина затрат на оплату труда собственного персонала $Z(A,a_i)$ рассчитывается по следующей формуле (рис. 3, уровень 6):

$$Z(A, a_i) = Z(A) + Z(a_i). \quad (11)$$

Совокупная величина затрат на оплату труда собственного персонала и на эксплуатацию собственной техники $Z(A,a_i,b_i)$ рассчитывается по следующей формуле (рис. 3, уровень 7):

$$Z(A, a_i, b_i) = Z(A, a_i) + Z(b_i). \quad (12)$$

Совокупная величина затрат на оплату труда подрядной организации и на эксплуатацию арендованной техники $Z(a_j,b_j)$ рассчитывается по следующей формуле (рис. 3, уровень 8):

$$Z(a_j, b_j) = Z(a_j) + Z(b_j). \quad (13)$$

Общая величина затрат Z рассчитывается по следующей формуле (рис. 3, уровень 9):

$$Z = Z(A, a_i, b_i) + Z(a_j, b_j). \quad (14)$$

Уравнения (1–14) определяют темпы потоков и уровни затрат, формируемых при эксплуатации ресурсов.

Результаты экспериментов имитационной модели системной динамики. Диаграмма потоков затрат (рис. 3) предназначена для структурирования затрат на оплату труда собственного персонала и персонала подрядной организации, а также суммы затрат, связанных с эксплуатацией и арендой техники, и реализована в пакете имитационного моделирования Anylogic (рис. 4).

Анимационная схема модели представлена в виде ГИС-карты, на которой ГИС-точками зеленого цвета отмечены энергообъекты. Их цвет меняется в зависимости от категории аварии: на красный цвет, если возникает авария 1 категории сложности; на оранжевый цвет, если возникает авария 2 категории сложности; на желтый цвет, если возникает авария 3 категории сложности. В течения всего времени устранения аварии возле ГИС-точки отображается текстовая информация о количестве задействованного собственного персонала и техники и отдельно о количестве привлеченного персонала и арендованной техники. По окончании времени устранения аварии ГИС-точка меняет цвет на зеленый, а текстовая информация исчезает (рис. 5).

Рисунок 4 – Схема модели системной динамики в Anylogic

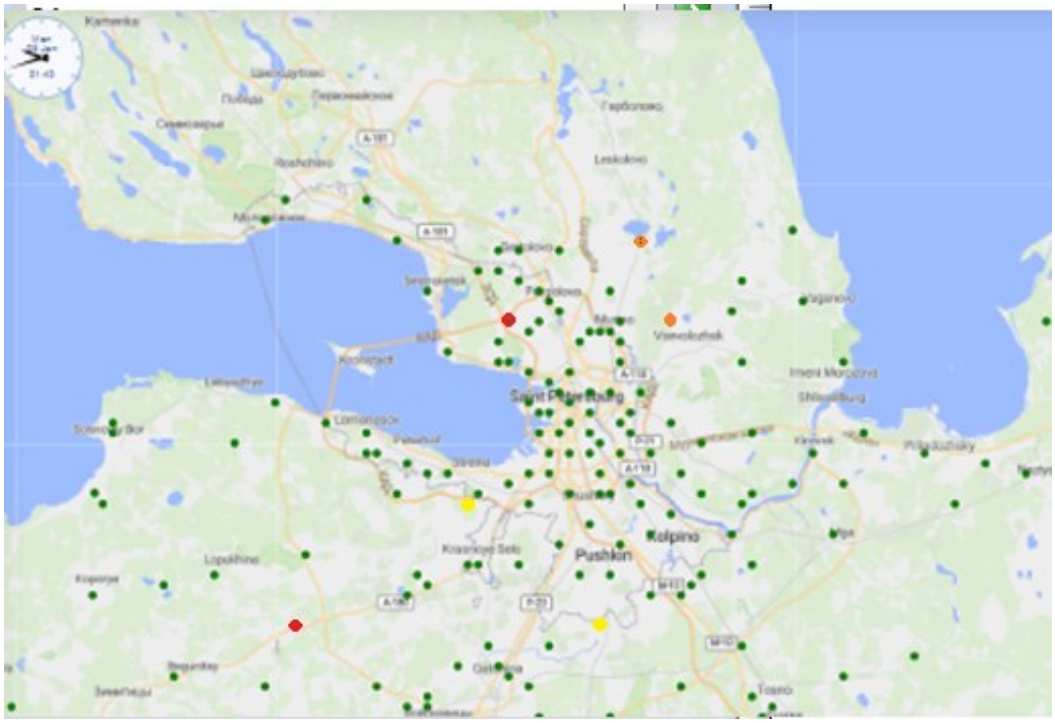


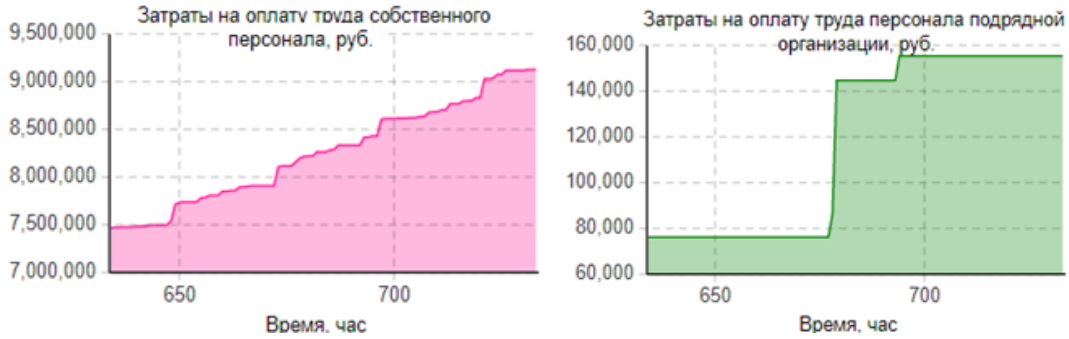
Рисунок 5 – ГИС-карта модели системной динамики в Anylogic

В зависимости от того, какую информацию необходимо получить о системе и какую задачу требуется решить с помощью имитационной модели, проводится планирование эксперимента, при этом должны быть учтены ограничения на ресурсы. Планирование машинных экспериментов, в сущности, представляет собой план получения необходимого объема информации, стоимость которого может варьироваться в зависимости от метода сбора и обработки данных. Эксперименты по компьютерному моделированию являются дорогостоящими с точки зрения времени и труда экспериментатора, а также затрат машинного времени. Чем больше усилий экспериментатор затрачивает на одно исследование, тем меньше времени он может потратить на другое, поэтому важно, чтобы он планировал экспериментирование с тем, чтобы получить как можно больше информации из каждого эксперимента. Таким образом, основной целью планирования экспериментов является определение наиболее важных факторов с тем, чтобы провести наименьшее количество прогонов в процессе экспериментирования с моделью.

Декомпозиция процесса планирования эксперимента с имитационной моделью анализа затрат процесса управления оперативно-выездными бригадами электросетевой компании представлена в нотации EPC (рис. 6). Предполагается, что работа с имитационной моделью будет выполняться Центром аналитических компетенций (ЦАК) – группой пользователей с ролью «Специалист ЦАК». Имитационная модель дает возможность пользователю «Специалист ЦАК» изменять значения параметров, установленных по умолчанию, перед запуском модели. Предполагается, что AnyLogic Private Cloud устанавливается внутри корпоративной сети электросетевой компании и пользователи ПК, подключенные к корпоративной сети (без ограничений по количеству пользователей), смогут просматривать модель, загруженную в AnyLogic Private Cloud пользователем «Специалист ЦАК».

Рисунок 6 – Этапы проведения эксперимента с моделью в нотации EPC

В окне запущенной модели на протяжении всего простого эксперимента наглядным образом аккумулируются затраты, связанные с использованием ресурсов для устранения аварий. Графики (рис. 7) предназначены для отображения временного тренда величины затрат на: оплату труда собственного персонала $Z(A,a_i)$; эксплуатацию собственной техники $Z(b_i)$; оплату труда подрядной организации $Z(a_j)$; эксплуатацию арендованной техники $Z(b_j)$.



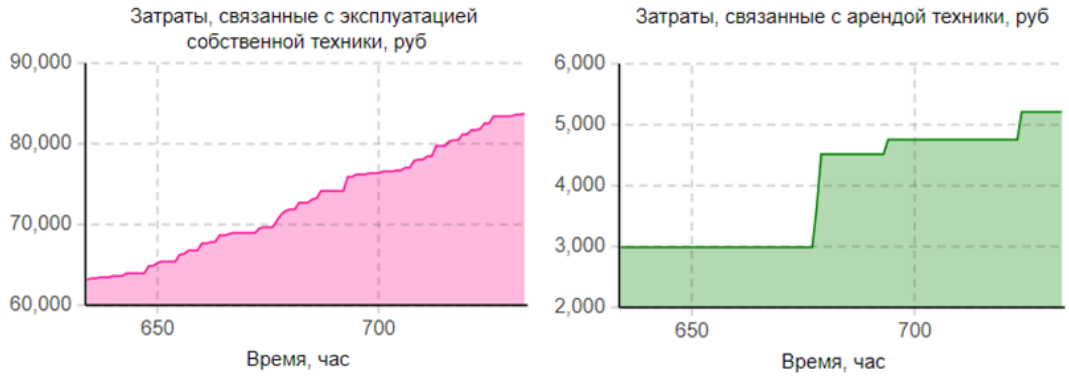
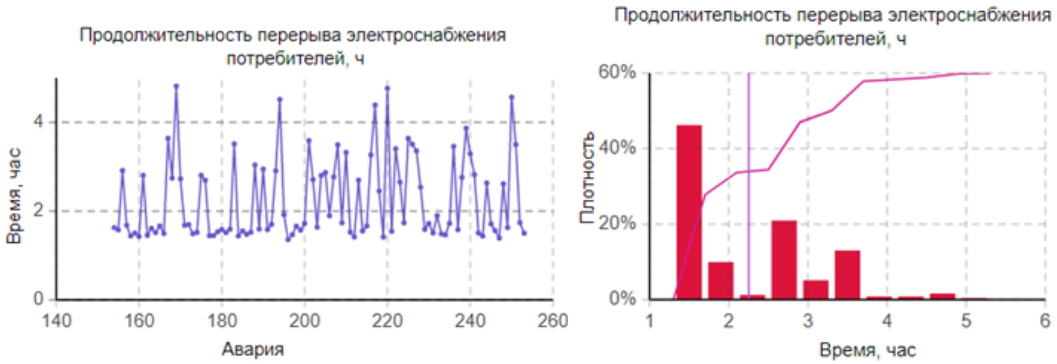


Рисунок 7 – Временные тренды величин затрат (фрагмент анимации)

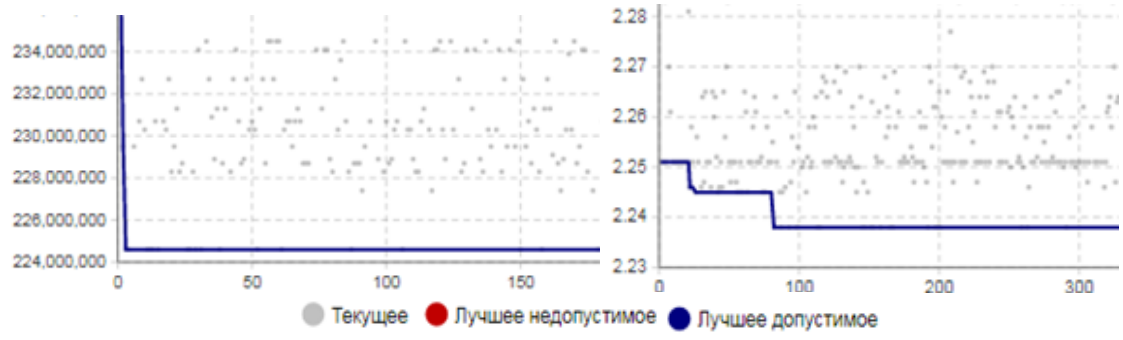
График «Продолжительность перерыва электроснабжения потребителей» (рис. 8*а*) предназначен для отображения временного тренда средней продолжительности перерыва электроснабжения потребителей. Гистограмма «Продолжительность перерыва электроснабжения» (рис. 8*б*) предназначена для отображения данных о продолжительности перерыва электроснабжения потребителей: гистограмма разбивает временную шкалу на 10 интервалов; тот интервал, в который попадает большее число продолжительностей, становится выше, что означает большую плотность попаданий в данный интервал.



а *б*

Рисунок 8 – Временные тренды продолжительности перерыва электроснабжения (фрагмент анимации)

При определении оптимальной численности дежурного персонала и оптимального количества техники с целью минимизации затрат (оптимизационный эксперимент 1) найден наилучший результат Z = 224 609 199,40 рублей, который достигается при A = 75 и B = 37 (рис. 9*а*). При определении оптимальной численности дежурного персонала и оптимального количества техники с целью сокращения продолжительности перерыва электроснабжения (оптимизационный эксперимент 2) найден наилучший результат (продолжительность прекращения электропередачи, равная 2,238 ч), который достигается при значениях A = 111 и B = 50 (рис. 9*б*).



а *б*

Рисунок 9 – Эксперименты с моделью

Имитационная модель системной динамики была разработана для ПАО «Ленэнерго» и является базовой моделью для создания моделей аналогичных распределительных электросетевых компаний. Для адаптации системы к условиям деятельности других электросетевых компаний имеется возможность добавления и (или) изменения значимых факторов, влияющих на уровень абстракции, настройки алгоритма формирования оперативно-выездными бригадами, задания мест расположений энергообъектов на ГИС-карте и др.

Заключение. Проведена формализация задачи системного моделирования динамики затрат процесса управления оперативно-выездными бригадами электросетевой компании. Рассмотрен детальный пример формирования затрат для поочередно возникающих 5 аварий и технологических нарушений в электросетевой компании. Показана динамика изменения структуры затрат в виде гистограммы соотношения затрат на собственные и привлеченные ресурсы. Построены детальные временные графики процесса управления затратами при возникновении аварий. Формализованы уравнения, определяющие темпы потоков и уровни затрат, формируемых при эксплуатации собственных и привлеченных ресурсов. Разработана концептуальная структура модели системной динамики на основе диаграммы потоков Форрестера. Реализована имитационная модель с элементами системной динамики, наложенная на ГИС-карту, для определения оптимального количества ресурсов, задействованных при устранении аварий, при которых средняя длительность перерыва электроснабжения потребителей является минимальной, а также для оптимизации затрат, связанных с оплатой труда персонала и с эксплуатацией и арендой техники, задействованных при устранении аварий. Составлен план экспериментов для системной модели динамики затрат при устранении аварий в электросетевом комплексе, проведены серии экспериментов, выявлены оптимальные характеристики по показателям численности персонала и общего количества задействованной техники.

Библиографический список

1. Zio, E. Challenges in the vulnerability and risk analysis of critical infrastructures / E. Zio // Reliab. Eng. Syst. Saf. – 2016. – Vol. 152. – P. 137–150. https://doi.org/10.1016/j.ress.2016.02.009.

2. Kröger ,W., Zio, E. Vulnerable systems / W. Kröger , E. Zio. – London : Springer; 2011.

3. Концепция обеспечения надежности в электроэнергетике : моногр. / Н. И. Воропай и др. – Москва : Энергия, 2013. – 212 с.

4. Концепция «Цифровая трансформация 2030». – Москва : ПАО «Россети», 2018. – Режим доступа: https://www.rosseti.ru/investment/Kontseptsiya_Tsifrovaya_transformatsiya_2030.pdf, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 06.05.2021).

5. Орлов, П. С. Технические мероприятия повышения надежности электроснабжения в АПК / П. С. Орлов, В. В. Морозов, С. П. Кочкин // Вестник АПК Верхневолжья. – 2017. – № 3 (39). – С. 94–100.

6. Проталинский, О. М. Система оптимального управления производственными активами энергетических предприятий / О. М. Проталинский, И. О. Проталинский, О. Н. Кладов // Автоматизация и IT в энергетике. – 2017. – № 4 (93). – С. 5–8.

7. Ханова, А. А. Управление затратами грузового порта на основе функционально-стоимостного анализа / А. А. Ханова, А. С. Пономарева // Известия высших учебных заведений. Северо-Кавказский регион. Технические науки. – 2011. – № 3 (161). – С. 116–119.

8. Проталинский, О. М. Адаптивная система прогнозирования надежности технологического оборудования объектов энергетики / О. М. Проталинский, И. А. Щербатов, А. А. Ханова, И. О. Проталинский // Информатика и системы управления. – 2019. – № 1 (59). – С. 93–105.

9. Бородин, В. А. Оптимизация ремонтных программ энергетического оборудования с использованием методов искусственного интеллекта / В. А. Бородин, А. В. Андрюшин, О. М. Проталинский, А. А. Ханова // Управление развитием крупномасштабных систем MLSD'2019 : материалы Двенадцатой международной конференции / под общ. ред. С. Н. Васильева, А. Д. Цвиркуна. – Москва : ИПУ РАН, 2019. – С. 576–579.

10. Горбунов, И. Н. Применение нейронных сетей в целях определения места повреждения воздушных и кабельных линий электропередачи / И. Н. Горбунов, С. Г. Захаренко, С. А. Захаров, Т. Ф. Малахова // Горное оборудование и электромеханика. – 2019. – № 4 (144). – С. 48–55.

11. Tanaka, M. Application of Kohonen's self-organizing network to the diagnosis system for rotating machinery / M. Tanaka, M. Sakawa, I. Shiromaru, T. Matsumoto // Proc. IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century. – 1995. – Vol. 5. – P. 4039–4044.

12. Малафеев, А. В. Использование теории нечетких множеств для оценки производственных рисков при управлении режимами промышленной системы электроснабжения / А. В. Малафеев, А. И. Юлдашева // Электроэнергетика глазами молодежи : труды VI Международной научно-технической конференции. – 2015. – С. 294–297.

13. Бобырь, М. В. Анализ методов повышения надежности нечетких систем / М. В. Бобырь, Н. А. Милостная // Вестник компьютерных и информационных технологий. – 2017. – № 7 (157). – С. 22–30.

14. Степанов, В. М. Имитационное и физическое моделирование систем электроснабжения для повышения надежности их работы / В. М. Степанов, И. М. Базыль // Известия Тульского государственного университета. Технические науки. – 2015. – № 12–2. – С. 139–142.

15. Дробов, А. В. Имитационная модель надежности системы электроснабжения / А. В. Дробов, В. Н. Галушко // Повышение надежности и энергоэффективности электротехнических систем и комплексов : межвузовский сборник научных трудов (с международным участием). – 2016. – С. 228–231.

16. Уразалиев, Н. С. Концептуальная структура имитационной модели логистических процессов управления ремонтами предприятия электрических сетей / Н. С. Уразалиев, А. А. Ханова, В. С. Тумпуров // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. – 2018. – № 2. – С. 91–100.

17. Борщев, А. Как строить простые, красивые и полезные модели сложных систем / А. Борщев // Имитационное моделирование. Теория и практика. ИММОД : материалы конференции. – Казань : Фэн АН РТ, 2013. – Т. 1. – С. 21–34.

18. Лычкина, Н. Н. Ретроспектива и перспектива системной динамики: анализ динамики развития / Н. Н. Лычкина // Бизнес-информатика. – 2009. – № 3 (9). – С. 55–67.

19. Маслобоев, А. В. Имитационное моделирование развития инновационных процессов на основе метода системной динамики и агентных технологий / А. В. Маслобоев // Качество. Инновации. Образование. – 2009. – № 3 (46). – С. 34–43.

20. Дворянчиков, А. Я., Редько, С. Г. Применение системной динамики для анализа поведения системы управления обучением / А. Я. Дворянчиков, С. Г. Редько // Инновации. – 2017. – № 3 (221). – С. 95–101.

21. Калажоков, Х. Х. Системное моделирование динамики сложных систем и процессов применительно к динамике развития региона / Х. Х. Калажоков, Ф. Х. Увижева // Известия Кабардино-Балкарского научного центра РАН. – 2017. – № 4 (78). – С. 5–10.

22. Форрестер, Джей. Основы кибернетики предприятия (Индустриальная динамика) / Форрестер Джей. – Москва : Прогресс, 1971. – 340 с.

References

1. Zio, E. Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliab. Eng. Syst. Saf.*, 2016, vol. 152, pp. 137–50.

2. Kröger, W., Zio, E. *Vulnerable systems*. London, Springer, 2011.

3. Kontseptsiya obespecheniya nadezhnosti v elektroenergetike [The concept of ensuring reliability in the electric power industry]. *Energiya* [Energy], 2013. 212 p.

4. *Kontseptsiya «Tsifrovaya transformatsiya 2030»* [Concept “Digital Transformation 2030”]. Moscow, “Rosseti” PJSC, 2018. Available at: https://www.rosseti.ru/investment/Kontseptsiya_Tsifrovaya _transformatsiya_2030.pdf (accessed 06.05.2021).

5. Orlov, P. S., Morozov, V. V., Kochkin, S. P. Tekhnicheskie meropriyatiya povysheniya nadezhnosti elektrosnabzheniya v APK [Technical measures to improve the reliability of power supply in the agro-industrial complex]. *Vestnik APK Verhnevolzhya* [Bulletin of the Agro-Industrial Complex of the Upper Volga Region], 2017, no. 3 (39), pp. 94–100.

6. Protalinskiy, O. M., Protalinskiy, I. O., Kladov, O. N. Sistema optimalnogo upravleniya proizvodstvennymi aktivami energeticheskikh predpriyatiy [System of optimal management of production assets of energy enterprises]. *Avtomatizatsiya i IT v energetike* [Automation and IT in Energy], 2017, no. 4 (93), pp. 5–8.

7. Khanova, A. A., Ponomareva, A. S. Upravlenie zatratami gruzovogo porta na osnove funktsionalno-stoimostnogo analiza [Management of freight port costs on the basis of functional-tax-value analysis]. *Izvestiya vysshikh uchebnykh zavedeniy. Severo-Kavkazskiy region. Tekhnicheskie nauki* [News of higher educational institutions. North Caucasus region. Technical Sciences], 2011, no. 3 (161), pp. 116–119.

8. Protalinskiy O. M., Shcherbatov I. A., Khanova A. A., Protalinskiy I. O. Adaptivnaya sistema prognozirovaniya nadezhnosti tekhnologicheskogo oborudovaniya obektov energetiki [Adaptive system for predicting the reliability of technological equipment of energy facilities]. *Informatika i sistemy upravleniya* [Informatics and Control Systems], 2019, no. 1 (59), pp. 93–105.

9. Borodin, V. A., Andryushin, A. V., Protalinsky, O. M., Khanova, A. A. Optimizatsiya remontnykh programm energeticheskogo oborudovaniya s ispolzovaniem metodov iskusstvennogo intellekta [Optimization of repair programs of energy equipment using artificial intelligence methods]. *Upravlenie razvitiem krupnomasshtabnykh sistem MLSD'2019 : materialy Dvenadtsatoy mezhdunarodnoy konferentsii* [Management of the development of large-scale systems MLSD'2019 : proceedings of the Twelfth International Conference], 2019, pp. 576–579.

10. Gorbunov, I. N., Zakharenko, S. G., Zaharov, S. A., Malakhova, T. F. Primenenie neyronnykh setey v tselyakh opredeleniya mesta povrezhdeniya vozdushnykh i kabelnykh liniy elektroperedachi [Use of neural networks to determine the location of damage to overhead and cable power lines]. *Gornoe oborudovanie i elektromekhanika* [Mining Equipment and Electromechanics], 2019, no. 4 (144), pp. 48–55.

11. Tanaka, M., Sakawa, M., Shiromaru, I., Matsumoto, T. Application of Kohonen's self-organizing network to the diagnosis system for rotating machinery. *Proc. IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, 1995, vol. 5, pp. 4039–4044.

12. Malafeev, A. V., Yuldasheva, A. I. Ispolzovanie teorii nechetkikh mnozhestv dlya otsenki proizvodstvennykh riskov pri upravlenii rezhimami promyshlennoy sistemy elektrosnabzheniya [Using the theory of fuzzy sets to assess production risks in managing the modes of an industrial power supply system]. *Elektroenergetika glazami molodezhi : trudy VI Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii* [Electricity industry through the eyes of young people : proceedings of the VI International Scientific and Technical Conference], 2015, pp. 294–297.

13. Bobyr, M. V., Milostnaya, N. A. Analiz metodov povysheniya nadezhnosti nechetkikh sistem [Analysis of methods for improving the reliability of fuzzy systems]. *Vestnik kompyuternykh i informatsionnykh tekhnologiy* [Bulletin of Computer and Information Technologies], 2017, no. 7 (157), pp. 22–30.

14. Stepanov, V. M., Bazyl, I. M. Imitatsionnoe i fizicheskoe modelirovanie sistem elektro-snabzheniya dlya povysheniya nadezhnosti ikh raboty [Simulation and physical modeling of electrical supply systems to increase the reliability of their work] *Izvestiya Tulskogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [Izvestia of the Tula State University. Technical Sciences], 2015, no. 12–2, pp. 139–142.

15. Drobov, A. V., Galushko, V. N. Imitatsionnaya model nadezhnosti sistemy elektrosnabzheniya [Simulation model of reliability of the power supply system]. *Povyshenie nadezhnosti i energoeffektivnosti elektrotekhnicheskikh sistem i kompleksov : mezhvuzovskiy sbornik nauchnykh trudov* [Improving the reliability and energy efficiency of electrical systems and systems : inter-university collection of scientific works], 2016, pp. 228–231.

16. Urazaliev, N. S., Khanova, A. A., Tumpurov, V. S. Konceptualnaya struktura imitatsionnoy modeli logisticheskikh protsessov upravleniya remontami predpriyatiya elektricheskikh setey [Conceptual structure of the imitation ministry of logistics processes for managing repairs of the electric grid enterprise]. *Vestnik Astrahanskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Upravlenie, vychislitelnaya tekhnika i informatika* [Bulletin of Astrakhan State Technical University. Series: Management, Computing and Computer Science], 2018, no. 2, pp. 91–100.

17. Borshchev, A. Kak stroit prostye, krasivye i poleznye modeli slozhnyhk sistem [How to build simple, beautiful and useful models of complex systems]. *Imitatsionnoe modelirovanie. Teoriya i Praktika. IMMOD : materialy konferentsii* [Simulation Simulation. Theory and Practice. IMMOD : materials of conference]. Kazan, Fen AN RT Publ., 2013, vol. 1, pp. 21–34.

18. Lychkina, N. N. Retrospektiva i perspektiva sistemnoy dinamiki: analiz dinamiki razvitiya [Retrospective and perspective of system dynamics. Development Dynamics Analysis]. *Biznes-informatika* [Business Informatics], 2009, no. 3 (9), pp. 55–67.

19. Masloboev, A. V. Imitatsionnoe modelirovanie razvitiya innovatsionnykh protsessov na osnove metoda sistemnoy dinamiki i agentnykh tekhnologiy [Simulation modeling of the development of innovative processes based on the method of system dynamics and agent technologies]. *Kachestvo. Innovatsii. Obrazovanie* [Quality. Innovation. Education], 2009, no. 3 (46), pp. 34–43.

20. Dvoryanchikov, A. Ya., Redko, S. G. Primenenie sistemnoy dinamiki dlya analiza povedeniya sistemy upravleniya obucheniem [Application of system dynamics for analysis of behavior of the training management system]. *Innovatsii* [Innovations], 2017, no. 3 (221), pp. 95–101.

21. Kalazhokov, H. H., Uvizheva, F. H. Sistemnoe modelirovanie dinamiki slozhnykh sistem i protsessov primenitelno k dinamike razvitiya regiona [Systematic simulation of the dynamics of complex systems and processes in relation to the dynamics of the development of the region]. *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [Izvestia of the Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences], 2017, no. 4 (78), pp. 5–10.

22. Forrester, Dzhey. *Osnovy kibernetiki predpriyatiya (Industrialnaya dinamika)* [Bases of enterprise cybernetics (Industrial dynamics)]. Moscow, Progress Publ., 1971. 340 p.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

УДК 004.89

НЕЙРОНЕЧЕТКАЯ МОДЕЛЬ И ПРОГРАММНЫЙ КОМПЛЕКС ФОРМИРОВАНИЯ БАЗ ЗНАНИЙ ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ОБЪЕКТОВ

Статья поступила в редакцию 18.02.2022, в окончательном варианте – 24.02.2022.

Катасёва Дина Владимировна[1], Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ, 420111, Российская Федерация, г. Казань, ул. К. Маркса, 10,
аспирант, ORCID: 0000-0001-6141-8329, e-mail: DVKataseva@kai.ru

Описана задача оценки состояния объектов. Для повышения эффективности ее решения и снижения влияния человеческого фактора актуализирована целесообразность использования систем поддержки принятия решений и формирования баз знаний нечетко-продукционного типа. В качестве инструмента их формирования предложено использовать нейронечеткую модель на основе обучения нечеткой нейронной сети. Рассмотрены особенности исходных данных для обучения. Предложен вид нечетко-продукционных правил, составляющих базу знаний. Рассмотрены этапы разработки нейронечеткой модели, включающие инициализацию и настройку значений параметров нечеткой нейронной сети. Описан разработанный метод ее обучения, основанный на использовании генетического алгоритма. Для его реализации рассмотрены вопросы выбора и кодирования параметров обучения на примере треугольных функций принадлежности. Предложена фитнес-функция и критерий выбора лучшей хромосомы в генетическом алгоритме. На основе предложенного математического обеспечения разработан программный комплекс для формирования баз знаний. Представлена его структура, состав программных модулей и их функциональность. Приведены результаты апробации программного комплекса на примере формирования базы знаний для подбора геолого-технических мероприятий на нефтяном месторождении. Результаты апробации подтвердили эффективность предложенного подхода и возможность его использования при формировании баз знаний и построении систем поддержки принятия решений для оценки состояния объектов в различных предметных областях.

Ключевые слова: оценка состояния объекта, база знаний, нечеткая нейронная сеть, нейронечеткая модель, нечетко-продукционное правило, функция принадлежности, генетический алгоритм, система поддержки принятия решений, геолого-технические мероприятия, нефтяное месторождение

NEURO-FUZZY MODEL AND SOFTWARE COMPLEX FOR FORMING KNOWLEDGE BASES FOR OBJECTS STATE ASSESSING

The article was received by the editorial board on 18.02.2022, in the final version – 24.02.2022.

Kataseva Dina V., Kazan National Research Technical University named after A.N. Tupolev-KAI, 10 K. Marx St., Kazan, 420111, Russian Federation,
postgraduate student, ORCID: 0000-0001-6141-8329, e-mail: DVKataseva@kai.ru

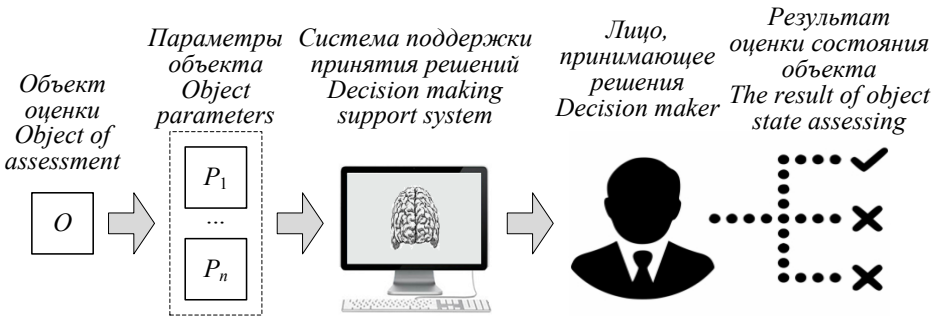
The task of objects state assessing is described. To improve the efficiency of its solution and reduce the human factor influence, the expediency of using decision-making support systems and the formation of fuzzy-production type knowledge bases has been updated. As a tool for their formation, it is proposed to use a neuro-fuzzy model based on a fuzzy neural network training. The features of the initial data for training are considered. A type of fuzzy-production rules that make up the knowledge base is proposed. The stages of neuro-fuzzy model construction, including the initialization and setting of the values of fuzzy neural network parameters, are considered. The developed method of its training based on the use of a genetic algorithm is described. For its implementation, the issues of choosing and coding learning parameters are considered on the example of triangular membership functions. A fitness function and a criterion for choosing the best chromosome in the genetic algorithm are proposed. Based on the proposed methods, a software package for the knowledge bases formation has been developed. Its structure, composition of program modules and their functionality are presented. The results of the software package approbation are presented on the example of the knowledge base formation for the selection of geological and technical measures in oil fields. The approbation results confirmed the effectiveness of the proposed approach and the possibility of its use in the knowledge bases formation and the decision-making support systems construction for objects state assessing.

Keywords: object state assessment, knowledge base, fuzzy neural network, neuro-fuzzy model, fuzzy production rule, membership function, genetic algorithm, decision support system, geological and technical measures, oil field

---

[1] Научный руководитель: Исмагилов Ильяс Идрисович, Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ, 420111, Российская Федерация, г. Казань, ул. К. Маркса, 10, доктор технических наук, профессор, ORCID: 0000-0002-0446-8204, e-mail: iiismag@mail.ru

Graphical annotation (Графическая аннотация)



Введение. В настоящее время в различных сферах человеческой деятельности возникает необходимость решения задач, связанных с оценкой состояния объектов [8]. Как правило, при решении таких задач ключевую роль играет человек – лицо, принимающее решения (ЛПР) [3]. На основе анализа значений параметров объекта P1,…, Pn он формирует результат оценки его состояния, выбирая оптимальное решение из множества допустимых. При этом ЛПР опирается на свое мнение, подкрепленное знаниями, опытом и интуицией.

Однако такая схема оценки сопряжена с рядом трудностей в силу наличия человеческого фактора. В качестве ЛПР может выступать как квалифицированный, так и неквалифицированный пользователь, ЛПР может ошибаться из-за невнимательности, усталости и т.д. Все это может приводить к ошибкам. Поэтому в настоящее время для повышения эффективности этого процесса актуально использование систем поддержки принятия решений (СППР) по оценке состояния объектов [1]. Рекомендации, поступающие от СППР, помогают ЛПР принять правильное решение, не ошибиться в выборе текущего состояния объекта.

Основными преимуществами применения СППР в задаче оценки состояния объекта являются [11]:

1) высокая скорость обработки информации при принятии решений;

2) адекватность рекомендуемых решений исходным данным, характеризующим оцениваемый объект;

3) снижение трудоемкости процесса принятия решений для человека.

Высокая скорость обработки информации достигается за счет автоматизации вычислений при анализе значений параметров объекта, выполнении процедуры логического вывода, сопоставлении входных данных с имеющимися в базе знаний системы правилами принятия решений. Адекватность рекомендуемых решений исходным данным об объекте оценки обусловлена тем, что интеллектуальная система всегда работает по одному алгоритму логического вывода и применяет для этого заранее сформированную базу знаний, адекватность которой обеспечивается в процессе ее формирования. Снижение трудоемкости выражается в уменьшении интеллектуальной нагрузки на ЛПР в процессе принятия решений.

Применение СППР главным образом приводит к снижению влияния человеческого фактора на процесс принятия решений и позволяет автоматизировать данный процесс.

При этом к СППР предъявляются следующие основные требования:

1) высокая точность оценки состояния объектов;

2) умение обосновать рекомендуемые решения в виде, пригодном для интерпретации ЛПР.

Высокая точность оценки состояния объектов должна достигаться за счет формирования адекватных баз знаний. Умение обосновывать рекомендуемые решения необходимо для их понимания и интерпретации человеком. Это, в свою очередь, повышает уровень доверия ЛПР к предлагаемым системой решениям, что позволяет сделать обоснованный выбор окончательного решения.

Одним из наиболее эффективных инструментов формирования баз знаний являются нечеткие нейронные сети (ННС) [2]. Рассмотрим особенности их использования.

Использование нечетких нейронных сетей для формирования баз знаний. Особенностью всех ННС является их гибридный характер [5]. Подобно классическим нейронным сетям они обучаются на исходных данных, формируя базу знаний нечетко-продукционного типа. Правила базы знаний могут быть использованы для решения поставленных задач, в данном случае – для задачи оценки состояния объектов.

На рисунке 1 представлена схема использования ННС для формирования базы знаний.

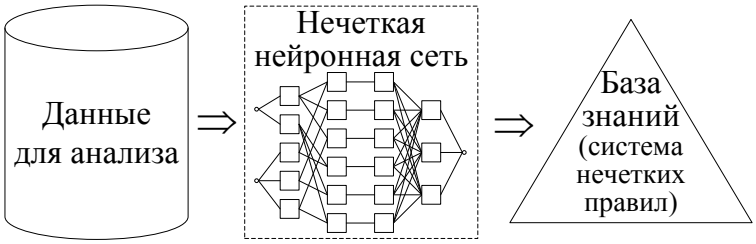


Рисунок 1 – Схема использования нечеткой нейронной сети для формирования базы знаний

Для реализации представленной схемы необходимо иметь набор данных, характеризующих решаемую задачу. Данные для анализа могут быть представлены в виде, как показано в таблице 1.

Таблица 1 – Общий вид таблицы данных для анализа

Входные переменные					**Выходная переменная**
1	2	3	…	*n*	

Таблица данных объемом N содержит n входных и одну выходную переменные, характеризующие оцениваемый объект. Каждая строка такой таблицы соответствует определенному объекту и содержит значения входных и выходной переменных. Значением выходной переменной является конкретное состояние объекта, в котором он находился при указанных условиях (конкретных значениях входных переменных). Следовательно, имея характеристики N объектов, можно сформировать исходную таблицу данных для анализа. При этом ННС выступает в качестве инструмента анализа данных [4].

Таким образом, для реализации описанной схемы формирования базы знаний на основе нейронечеткого подхода необходимо решить следующие основные задачи:

1) определить особенности исходных данных, характеризующих оцениваемые объекты и используемых для принятия решений;

2) предложить вид нечетких правил, учитывающих особенности исходных данных;

3) разработать нейронечеткую модель для формирования правил выбранного вида.

Особенности исходных данных. При принятии решений используются входные данные, характеризующие оцениваемые объекты. Объекты оценки могут быть различной природы. При этом данные, описывающие оцениваемые объекты, могут быть получены из различных источников. Поэтому для эффективного решения поставленной задачи необходимо понимать и учитывать их особенности, а именно возможную их разнотипность и некомплектность.

Разнотипность данных означает, что часть входных переменных в таблице для анализа является количественной и принимает непрерывные или дискретные значения, лежащие на числовой оси. Остальные входные переменные являются качественными и принимают категориальные (номинальные, порядковые или бинарные) значения. Некомплектность данных, используемых для принятия решений, означает ситуацию, когда в процессе оценки состояния объекта на входе СППР появляются данные с пропущенными значениями.

Вид нечетких правил. Для учета особенностей входных данных, используемых при оценке состояния объектов, предложен следующий вид нечетко-продукционных правил:

$$If\ x_1 = \vec{A}_1\ \&\ x_2 = \vec{A}_2\ \&\ \ldots\ x_i = \vec{A}_i\ \&\ \ldots\ x_n = \vec{A}_n\ Then\ y = B\,[CF], \tag{1}$$

где x_i (i=1..n) – входные переменные; $\vec{A}_i$ – четкие ($\vec{A}_i = A_i$) и нечеткие ($\vec{A}_i = \widetilde{A}_i$) ограничения на значения входных переменных, позволяющие обрабатывать, соответственно, качественные и количественные входные данные; A_i – четкие значения входных переменных; $\widetilde{A}_i = \{x_i, \mu_{\widetilde{A}_i}(x_i)\}$ – нечеткие значения входных переменных; y – выходная переменная; B – состояние объекта; $CF \in [0; 1]$ – вес правила.

Достоинства нечетко-продукционных правил вида (1) заключаются в следующем:

1) позволяют обрабатывать разнотипные входные данные;
2) поддаются однозначной лингвистической интерпретации;
3) характеризуются весом *CF* (для оценки состояния объекта).

Рассмотрим предложенную в работе нейронечеткую модель.

Разработка нейронечеткой модели для формирования баз знаний. Для построения любой интеллектуальной модели на основе методов машинного обучения требуется выполнить ряд последовательных этапов. Рассмотрим этапы построения нейронечеткой модели (рис. 2).

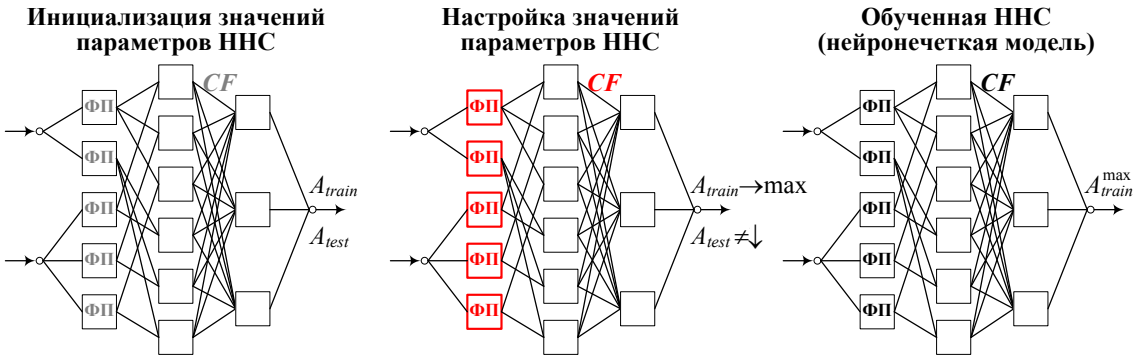


Рисунок 2 – Этапы построения нейронечеткой модели

Для построения нейронечеткой модели требуется выполнить этапы инициализации и настройки значений параметров ННС. На первом этапе необходимо определить начальные значения параметров функций принадлежности (ФП) и начальные значения весов CF. Также требуется рассчитать начальные значения точности классификации ННС на обучающей *(Atrain)* и тестовой *(Atest)* выборках данных.

На этапе настройки производится так называемый «тюнинг» модели [6], в результате которого определяются оптимальные значения указанных параметров с учетом следующего критерия оптимизации:

$$A_{train} = \frac{N_{train\_t}}{N_{train}} \to \max \mid A_{test} = \frac{N_{test\_t}}{N_{test}} \neq \downarrow, \tag{2}$$

где N_{train} – объем обучающей выборки; N_{train_t} – число правильно классифицированных обучающих данных ($N_{train\_t} \le N_{train}$); N_{test} – объем тестовой выборки; N_{test_t} – число правильно классифицированных тестовых данных ($N_{test\_t} \le N_{test}$); «$\neq\downarrow$» – условие неуменьшения точности классификации ННС на тестовой выборке данных.

Следовательно, при реализации этого этапа необходимо максимизировать точность классификации A_{train} нечеткой нейронной сети на обучающей выборке данных в условиях контроля ее переобучения (неуменьшения точности классификации ННС A_{test} на тестовой выборке).

Результатом выполнения указанных этапов является построенная нейронечеткая модель с настроенными значениями параметров. Для реализации отдельных этапов ее построения требуется разработка соответствующих методов и алгоритмов.

Инициализация является первичным этапом построения нейронечеткой модели. Рассмотрим его реализацию на примере треугольных ФП, используемых для формализации нечетких ограничений на значения входных количественных переменных в нечетких правилах. Выбор треугольных ФП обусловлен широтой и эффективностью их практического использования [10, 13], а также простотой задания и лингвистической интерпретацией человеком, что требуется для принятия обоснованных решений.

На рисунке 3 показаны примеры определения начальных значений параметров ФП.

При любом количестве выбираемых нечетких градаций используется равномерный метод гранулирования. Такой подход является классическим и позволяет максимально просто определить начальные значения параметров каждой функции принадлежности [15].

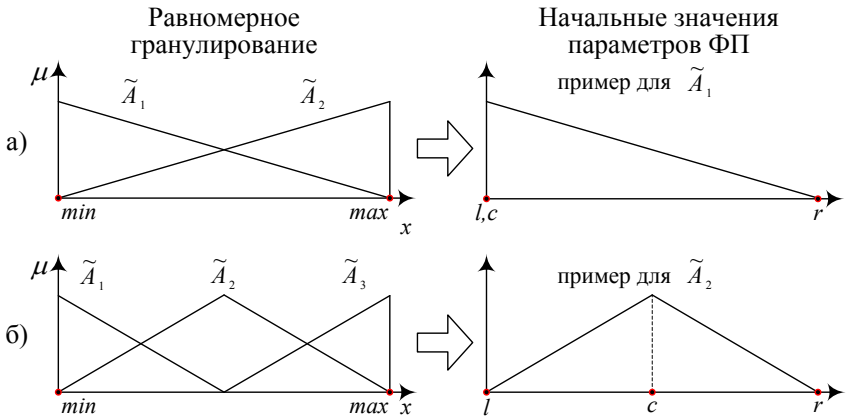


Рисунок 3 – Определение начальных значений параметров функций принадлежности

Так, например, в случае 2-х нечетких градаций задаются две треугольные ФП на отрезке [*min*; *max*], как показано на рисунке 3а. Границы этого отрезка соответствуют минимальному и максимальному значениям входного параметра *x* в обучающей выборке данных. Это позволяет определить начальные значения параметров ФП: *l* – левое основание, *c* – моду и *r* – правое основание. На рисунке 3а справа показан пример определения начальных значений параметров для первой функции принадлежности. Аналогично определяются начальные значения параметров ФП для 3-х (рис. 3б) и большего числа градаций.

После выполнения этого этапа необходимо обучить нечеткую нейронную сеть путем настройки значений ее параметров. Рассмотрим разработанные метод и алгоритмы, используемые для обучения ННС.

Существует большое количество методов оптимизации, используемых в машинном обучении при построении интеллектуальных моделей. Так, например, при обучении нейронных и нечетких нейронных сетей большое распространение получили градиентные методы [14], используемые в алгоритме обратного распространения ошибки, который стал классическим для построения таких моделей. Данный алгоритм предполагает вычисление и минимизацию среднеквадратичной ошибки выхода сети на основе вычисления градиентов по настраиваемым параметрам модели.

Однако в предложенной нечеткой нейронной сети функция ошибки принимает дискретные значения и зависит от дискретных параметров, поэтому применение понятия «градиент» для нее будет некорректным. Кроме того, градиент вычисляется для гладких функций. Используемые в работе треугольные функции принадлежности кусочно-линейны, т. е. не являются гладкими на всей области определения. Поэтому для обучения ННС предложено использовать генетический алгоритм (ГА) [9]. Такие алгоритмы получили широкое распространение. Они не требуют использования гладких ФП и могут быть реализованы с использованием различных эвристических приемов, связанных с кодированием параметров задачи, выполнением генетических операторов и выбором критериев для нахождения квазиоптимальных решений поставленных задач.

На рисунке 4 представлена схема разработанного метода обучения ННС с использованием генетического алгоритма.

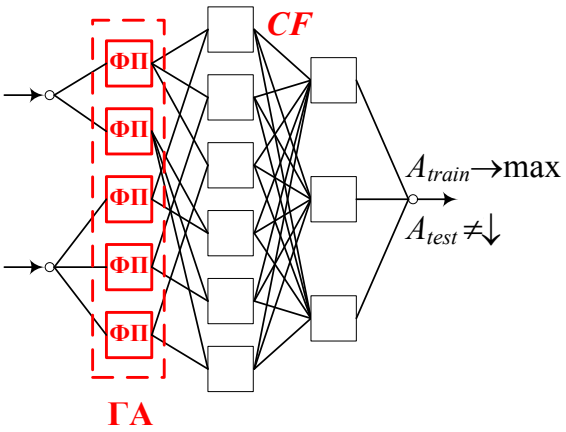


Рисунок 4 – Схема метода обучения нечеткой нейронной сети

Из рисунка видно, что ГА используется для настройки значений параметров ФП в ННС. При этом на каждой итерации генетической оптимизации происходит уточнение значений весовых коэффициентов *CF*. Целью обучения является максимизация A_{train} (точности классификации обучающих данных) при условии неуменьшения A_{test} (точности классификации тестовых данных).

Таким образом, для обучения ННС разработан метод итерационной генетической оптимизации, позволяющий настраивать значения параметров ФП и по результатам настройки определять итоговые значения весовых коэффициентов *CF*.

Рассмотрим вопросы выбора и кодирования параметров ФП, представленных на рисунке 5.

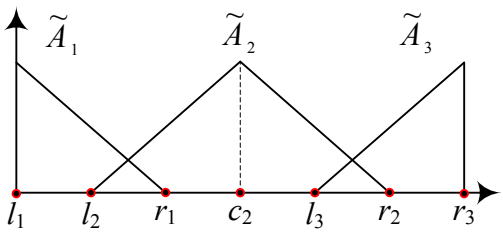


Рисунок 5 – Функции принадлежности и их параметры

На рисунке представлены три треугольные функции принадлежности:

- $\widetilde{A}_1$ с параметрами (l_1, r_1);
- $\widetilde{A}_2$ с параметрами (l_2, c_2, r_2);
- $\widetilde{A}_3$ с параметрами (l_3, r_3).

Следует отметить, что для задания крайних функций принадлежности (в данном случае это $\widetilde{A}_1$ и $\widetilde{A}_3$) достаточно иметь два параметра (в данном случае это l – левое основание и r – правое основание). Во всех остальных случаях ФП задаются тремя параметрами: l – левое основание, c – мода и r – правое основание.

Очевидно, что не все из представленных параметров $\{l_1, l_2, r_1, c_2, l_3, r_2, r_3\}$ являются настраиваемыми. Следовательно, необходимо выбирать такие параметры p, значения которых можно настраивать. Рассмотрим вопрос выбора настраиваемых параметров.

Обучение ННС заключается в настройке значений параметров ее функций принадлежности. При этом значения крайних параметров (в данном случае это параметры l_1 и r_3) не могут изменяться, иначе нарушится интерпретируемость соответствующих крайних функций принадлежности. Следовательно, $p \in \{l_2, r_1, c_2, l_3, r_2\}$. Значения именно этих параметров могут и должны изменяться.

Следующим вопросом является определение интервалов значений для настраиваемых параметров. Поскольку функции принадлежности должны быть интерпретируемы, упорядочены и не иметь точек разрыва, то в качестве указанных интервалов параметров $p \in (min; max)$ целесообразно выбрать следующие:

$$l_2 \in (l_1; r_1),\ r_1 \in (l_2; l_3),\ c_2 \in (l_2; r_2),\ l_3 \in (r_1; r_2),\ r_2 \in (l_3; r_3).$$

Далее необходимо закодировать значения настраиваемого параметра в заданном интервале. Для этого каждый интервал равномерно разобьем на $2^m + 1$ отрезков, где m – целое число. Тогда каждый параметр сможет принимать одно из 2^m значений:

$$H = \{H_1, \ldots, H_z, \ldots, H_Z\} = \{\overbrace{(0, 0, \ldots, 0)}^{m}, \overbrace{(0, 0, \ldots, 1)}^{m}, \ldots, \overbrace{(1, 1, \ldots, 1)}^{m}\},\ z = 1..Z,\ Z = 2^m,$$

где $H_1 = (0, 0, \ldots, 0)$ – хромосома, соответствующая наименьшему значению параметра; $H_Z = (1, 1, \ldots, 1)$ – хромосома, соответствующая его наибольшему значению; m – длина хромосомы (количество ее ген).

На рисунке 6 показан пример кодирования (для $m = 4$).

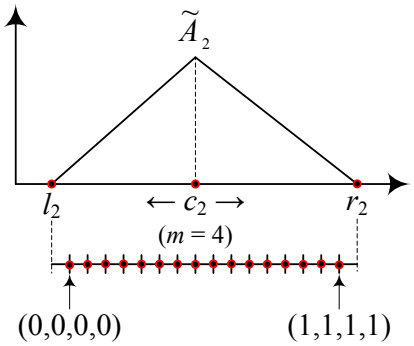


Рисунок 6 – Пример кодирования значений параметра функции принадлежности

Интервал (l_2; r_2) ограничивает возможные значения для настраиваемого параметра c_2. В данном случае этот параметр может принимать одно из 16 значений (так как $m = 4$, то $2^m = 2^4 = 16$).

Завершающим вопросом, требующим решения, является задание фитнес-функции для оценки значений настраиваемого параметра. Так как целью обучения ННС является максимизация точности ее классификации на обучающей выборке данных, то в качестве фитнес-функции целесообразно задать следующую:

$$F(H_z) = \frac{N_{train\_t}}{N_{train}}, \ (z = 1 \ldots Z, Z = 2^m). \tag{3}$$

Лучшая хромосома максимизирует точность классификации ННС на обучающей выборке (в соответствии с (3)) при условии неуменьшения ее точности классификации A_{test} на тестовой выборке:

$$F(H_z) \rightarrow \max_{\forall H_z} | A_{test}(H_z) \neq\downarrow, \ \ A_{test}(H_z) = \frac{N_{test\_t}}{N_{test}}. \tag{4}$$

Таким образом, выбор настраиваемых параметров функций принадлежности, определение интервалов и кодирование их возможных значений, а также задание фитнес-функции с критерием выбора лучшей хромосомы позволили реализовать классический ГА [7] для построения нейронечеткой модели.

Программный комплекс формирования баз знаний. Описанные модель и алгоритмы реализованы в программном комплексе, создание которого осуществлялось на базе интегрированной среды разработки Visual Studio 2019 на языке C#. Программный комплекс имеет модульную структуру (рис. 7).

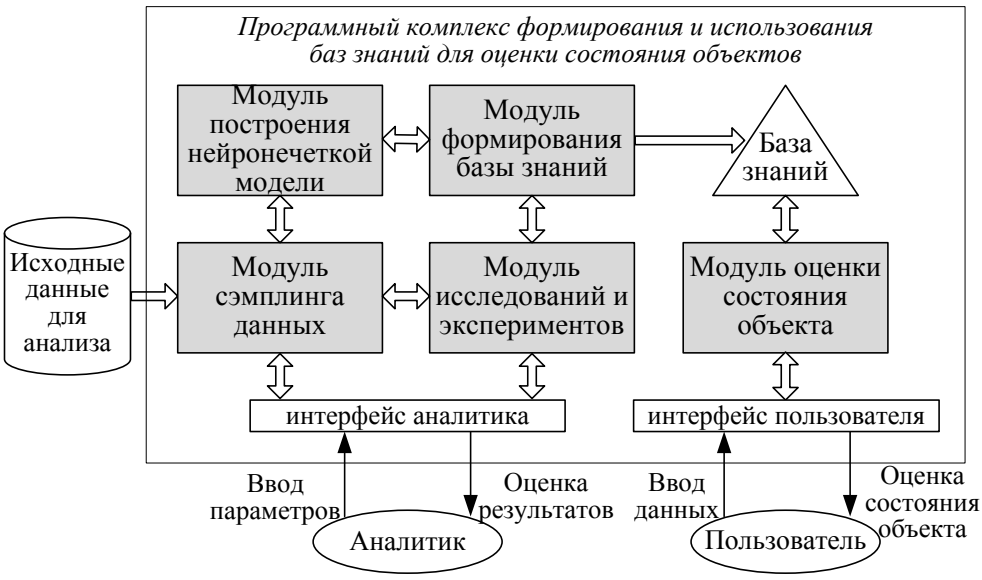


Рисунок 7 – Структура программного комплекса

Программный комплекс состоит из следующих модулей:

1) модуль сэмплинга данных (производит загрузку исходных данных, их визуализацию, формирование выборок данных для анализа);

2) модуль построения нейронечеткой модели (строит и визуализирует структуру ННС, инициализирует значения параметров ННС, производит ее обучение);

3) модуль формирования базы знаний (формирует систему правил на основе нейронечеткой модели, визуализирует правила базы знаний);

4) модуль оценки состояния объектов (производит оценку состояния объектов, визуализирует результаты оценки);

5) модуль исследований и экспериментов (позволяет производить исследования нейронечеткой модели и методов ее построения, визуализирует результаты исследований).

Каждый программный модуль обладает соответствующими функциональными возможностями, в совокупности составляющими возможности всего программного комплекса. Кроме того, программный комплекс имеет интерфейсы для работы с ним аналитика и конечного пользователя. Под управлением аналитика производится формирование базы знаний. Пользователь использует сформированную базу знаний непосредственно для оценки состояния объектов.

Следует отметить, что программный комплекс прошел апробацию при формировании баз знаний в различных предметных областях, в частности в нефтяной отрасли. Рассмотрим результаты его апробации.

Апробация программного комплекса в нефтяной отрасли. Одной из актуальных задач в нефтяной отрасли является подбор геолого-технических мероприятий (ввод скважин в эксплуатацию) на нефтяном месторождении. Для обучения ННС и формирования базы знаний использовались исходные данные, описываемые следующими параметрами:

1) общие сведения о каждой скважине:
- порядковый номер скважины;
- промысловый номер добывающей скважины;
- год начала отбора и др.;

2) геолого-геофизические данные нефтяного пласта:
- проницаемость абсолютная;
- пористость;
- нефтенасыщенность начальная и др.;

3) параметры на год начала отбора:
- площадь зоны дренирования;
- начальные балансовые запасы нефти;
- текущие балансовые запасы нефти и др.

Исходные данные были представлены по 46 объектам нефтяного месторождения с карбонатными коллекторами. Объем данных составил 3398 записей.

В таблице 2 представлен фрагмент исходных данных для анализа.

В данной таблице в качестве входных параметров указаны следующие:
- pronic – абсолютная проницаемость породы;
- poris – пористость породы;
- nasich – начальная нефтенасыщенность;
- tolsh – начальная нефтенасыщенная толщина;
- plosh – площадь зоны дренирования;
- bal1 – начальные балансовые запасы нефти;
- bal2 – текущие балансовые запасы нефти;
- kohvat – коэффициент охвата заводнением;
- rnagn – расстояние до ближайшей нагнетательной скважины;
- debsos – средний дебит соседних скважин;
- obvod – минимальная обводненность соседних скважин.

Единственный выходной параметр «result» характеризует описываемую скважину как эффективную в применении (result = 1 при дебите скважины > 5) или неэффективную (result = 0 при дебите скважины ≤ 5).

Таблица 2 – Фрагмент исходных данных по скважинам

pronic	poris	nasich	tolsh	plosh	bal1	bal2	kohvat	rnagn	debsos	obvod	result
56,7	0,131	0,685	5,8	12	22	20,7	0,1375	7586	0,2	0,5	0
12	0,119	0,656	4,4	12	14,6	11,8	0,3531	17111	0,1	0,597	0
56,7	0,131	0,685	1,6	0,1	0	0	0,3543	0	0	0	0
56,7	0,131	0,685	6,4	12	24,3	23,8	0,0418	12390	0,4	0,794	0
56,7	0,131	0,685	8,4	4,4	5,8	5,8	0,005	5798	1,4	0,095	1
87	0,162	0,695	2,8	12	13,3	11,6	0,2483	12366	0	0	0
56,7	0,131	0,685	4,2	12	15,9	15	0,1424	11272	0	0	0
54,8	0,133	0,683	7,1	12	21	19,8	0,13	7855	0,5	0,955	0
30	0,148	0,49	1,4	12	4,3	4	0,2084	1005	0	0	0
10,4	0,108	0,644	5,2	12	11,3	10,5	0,3722	13787	0,3	0,773	0
50,7	0,138	0,714	9,9	12	34,4	31,1	0,0925	4270	0	0	0
56,7	0,131	0,685	4,6	7,3	5,3	4,9	0,1638	3190	0,9	0,837	0
45,3	0,12	0,681	14,9	12	32,3	31,4	0,0719	34181	1	0,265	0
2	0,097	0,796	16,1	12,4	34,8	29,4	0,4396	25489	0	0,065	0
3	0,116	0,749	16,9	12	60,4	59,9	0,0136	4034	1,1	0,709	0
13,2	0,121	0,687	3,7	12	14,8	14,3	0,0616	0	0	0	0
1	0,096	0,565	2,8	11,3	8,1	7,6	0,1563	0	0,3	0,755	0
6	0,127	0,66	19,7	0,4	1,4	1,4	0,0284	0	9,5	0,102	0
5,8	0,111	0,792	6,5	12	24,5	24,3	0,0097	0	1,4	0,215	1
4	0,123	0,6	5,8	12	24,3	23,7	0,0469	0	0	0	0
13,9	0,13	0,77	4,7	12	26,7	25,8	0,0529	0	0,9	0,794	1
13,9	0,094	0,74	5,2	13,1	28,5	26,1	0,14	0	0,2	0,778	0
1,3	0,106	0,623	5,8	12	16	14,9	0,243	0	0,2	0,858	0

Перед формированием базы знаний произведена подготовка исходных данных к анализу. Никакой метод машинного обучения не даст положительного результата, если данные будут плохого качества. Поэтому важным является оценка качества исходных данных и их подготовка к анализу.

В наборе исходных данных встречались выбросы и аномалии. При подготовке исходных данных к анализу выполнен их поиск и корректировка, поскольку они являются одним из факторов, снижающих качество данных и достоверность результатов анализа. Алгоритм поиска выбросов и аномалий позволил оценить качество исходных данных и выработать рекомендации по его улучшению. Всего в исходных данных объемом 3398 строк было найдено 917 строк с выбросами и 211 строк с аномалиями. Параметр «tip» оказался непригодным для анализа. Единственный параметр «obvod» оказался полностью пригодным к анализу. К остальным значениям параметров были применены методы повышения качества исходных данных [12]: удаления записей с аномальными значениями и заменой выбросов на медианное значение. После предобработки объем исходных данных сократился на 211 строк (с 3398 до 3187) в соответствии с количеством удаленных записей с аномальными значениями.

Следующим этапом подготовки данных явился корреляционный анализ между параметрами. При построении интеллектуальной модели целесообразно оценить зависимости в исходных данных для анализа, а именно выбрать систему входных признаков, значимо влияющих на выходной результат. Данная задача актуальна при обучении ННС, способных эффективно обучаться в пространстве небольшой (до 10–12 входных параметров) размерности.

В результате корреляционного анализа данных произведен выбор входных параметров. В таблице 3 представлены значения рассчитанных коэффициентов корреляции.

Из таблицы видно, что не все входные параметры имеют корреляционную зависимость с выходным. Так, например, такие параметры, как «проницаемость», «пористость», «средний дебит соседних скважин» и «минимальная обводненность соседних скважин» имеют относительно высокую корреляцию с выходным параметром «результат». Это означает, что указанные параметры целесообразно использовать в обучении. При этом порог корреляции, определяющий количество и состав входных параметров, подбирался экспериментально.

После подготовки исходных данных и их анализа произведено формирование баз знаний с различным набором комбинаций параметров. Проведены следующие исследования:

- оценка влияния порога корреляции в исходных данных на результаты моделирования;
- оценка влияния порога эффективности скважины на результаты обучения ННС;
- оценка влияния вариантов расчета выходной переменной (первый ненулевой дебит скважины за последние 5 лет, средний дебит скважины за первые 2 года эксплуатации, средний дебит скважины за первые 3 года эксплуатации, максимальный дебит скважины за последние 5 лет) на результаты обучения ННС.

Таблица 3 – Результаты корреляционного анализа

Входные параметры	Коэффициент корреляции с выходным параметром
gru	–0,03
pronic	–0,153
poris	–0,136
nasich	0,015
tolsh	0,02
plosh	0,028
bal1	0,01
bal2	0,007
podv1	0,017
podv2	0,017
kohvat	0,012
rnagn	0,013
debsos	0,19
obvod	–0,189

Анализ проведенных экспериментов позволил выявить ряд особенностей и закономерностей. Во-первых, при уменьшении порога корреляции с 0,15 до 0,05 и, соответственно, при увеличении числа входных параметров в модели точность классификации данных незначительно возрастает (с 90 до 90,53 % на обучающей выборке данных). Однако при этом существенно возрастает время построения модели (со 140 до 893 минут). Следовательно, не имеет смысла включать в модель большое число входных параметров. Для эффективной классификации данных достаточно использования четырех параметров.

Во-вторых, при изменении порога эффективности скважины относительно начального значения точность классификации незначительно снижается. Следовательно, экспертная оценка эффективности скважины на уровне 5 тонн в сутки является адекватной и согласуется с результатами экспериментов.

В-третьих, при подготовке данных к анализу лучшие результаты классификации достигаются при расчете среднего дебита нефти по скважине за первые 3 года эксплуатации. Соответственно, этот критерий наиболее адекватно (с точки зрения достигаемых результатов классификации) характеризует дебит нефти по скважине и, как следствие, оценку ее эффективности.

Таким образом, при формировании искомой базы знаний для подбора ГТМ (ввода скважин в эксплуатацию) на нефтяном месторождении реализованы следующие решения, отражающие оптимальный набор значений параметров:

1) число входных параметров: 4;

2) порог эффективности скважины: 5;

3) расчет эффективного дебита нефти (значения выходного параметра): средний дебит нефти по скважине за первые 3 года эксплуатации.

В результате анализа исходных данных и построения нейронечеткой модели сформирована база знаний для определения вариантов ГТМ (ввода скважин в эксплуатацию) на нефтяном месторождении. Рассмотрим фрагмент и характеристики сформированной базы знаний (табл. 4).

Таблица 4 – Фрагмент сформированной базы знаний

pronic	poris	debsos	obvod	result	CF
1	1	1	1	0	0,02
1	1	1	2	0	0,02
1	1	1	3	0	0,01
1	1	2	1	0	0,06
1	1	2	1	1	0,08
1	1	2	2	0	0,04
1	1	2	4	0	0,03
1	1	3	1	0	0,07
1	1	3	1	1	0,08
1	1	3	2	0	0,09
1	1	3	3	0	0,05
1	1	3	5	0	0,08

База знаний состоит из 448 нечетких правил. Каждое правило включает в себя следующие параметры:

- pronic – абсолютная проницаемость породы;
- poris – пористость породы;
- debsos – средний дебит соседних скважин;
- obvod – минимальная обводненность соседних скважин.

Значения входных параметров «1», «2», «3», «4», «5» означают, соответственно, «низкий», «ниже среднего», «средний», «выше среднего», «высокий».

Точность классификации данных при обучении ННС составила 97,54 %, точность классификации при тестировании составила 95,7 %, точность классификации при валидации модели – 95,68 %. Достигнутая точность классификации данных характеризует высокую аппроксимирующую способность сформированной базы знаний, а следовательно, ее адекватность и возможность практического использования.

Таким образом, полученные результаты точности классификации свидетельствуют о высокой обобщающей способности сформированной базы знаний.

Заключение. Описанные в данной работе нейронечеткая модель и программный комплекс прошли апробацию в управлении инвестиций и Центре обслуживания бизнеса ПАО «Татнефть», внедрены в бизнес-процесс формирования оптимальной программы геолого-технических мероприятий по всем объектам разработки в условиях ресурсных ограничений с учетом вариативности налоговых моделей. Использование результатов работы позволило:

- произвести выбор оптимальной стратегии проведения геолого-технических мероприятий по объектам разработки ПАО «Татнефть» в условиях ресурсных и финансовых ограничений;
- произвести ранжирование объектов разработки с целью рекомендации перехода на налоговую модель на дополнительный доход от добычи углеводородного сырья;
- снизить нагрузку на геолого-технологический персонал по рациональному выбору проектных скважин для бурения на нефтяных месторождениях.

Таким образом, описанный подход может быть использован при формировании баз знаний для оценки состояния объектов в различных предметных областях.

Библиографический список

1. Аждер, Т. Б. Системы поддержки принятия решений и информационные системы / Т. Б. Аждер, О. А. Гуреева // Уральский научный вестник. – 2019. – Т. 6, № 3. – С. 46–48.

2. Баринов, А. И. Использование модели нечетких нейронных сетей для формирования базы знаний по определению фишинговых сайтов / А. И. Баринов, Д. В. Катасёва, А. С. Катасёв // Вестник Технологического университета. – 2020. – Т. 23, № 10. – С. 64–67.

3. Башлыков, А. А. Роль человека-оператора как лица, принимающего решения, и элемента интеллектуальной системы управления сложными технологическими объектами / А. А. Башлыков // Автоматизация, телемеханизация и связь в нефтяной промышленности. – 2016. – № 12. – С. 10–18.

4. Глова, В. И. Преднастройка и оптимизация параметров нечеткой нейронной сети при формировании баз знаний экспертных систем / В. И. Глова, А. С. Катасёв, Г. С. Корнилов // Информационные технологии. – 2010. – № 5. – С. 15–19.

5. Гридина, Н. В. Построение гибридных нейронных сетей с использованием элементов нечеткой логики / Н. В. Гридина, И. А. Евдокимов, В. И. Солодовников // Искусственный интеллект и принятие решений. – 2019. – № 2. – С. 91–97.

6. Джуманов, О. И. Методы оптимизации идентификации нестационарных объектов на основе нейронечеткой сети с настройкой параметров вычислительных схем / О. И. Джуманов, С. М. Холмонов // Проблемы вычислительной и прикладной математики. – 2016. – № 3 (5). – С. 20–27.

7. Звонков, В. Б. Сравнительное исследование классических методов оптимизации и генетических алгоритмов / В. Б. Звонков, А. М. Попов // Вестник Сибирского государственного аэрокосмического университета им. академика М.Ф. Решетнева. – 2013. – № 4 (50). – С. 23–27.

8. Клюева, А. Р. Исследование и оценка состояния технических объектов на основе компьютерного моделирования / А. Р. Клюева // Инженерный вестник Дона. – 2018. – № 4 (51). – С. 120–132.

9. Комарцов, Л. Г. Исследование генетических алгоритмов для обучения многослойного персептрона / Л. Г. Комарцова, Д. С. Кадников // Нейрокомпьютеры: разработка, применение. – 2010. – № 12. – С. 12–19.

10. Коротеев, М. В. Формы функции принадлежности лингвистических переменных экономических показателей / М. В. Коротеев // Аудит и финансовый анализ. – 2012. – № 2. – С. 239–244.

11. Кравченко, Т. К. Создание систем поддержки принятия решений: интеграция преимуществ отдельных подходов / Т. К. Кравченко, Н. Н. Середенко // Искусственный интеллект и принятие решений. – 2012. – № 1. – С. 39–47.

12. Кузовлев, В. И. Повышение качества данных с использованием методики поиска аномалий на примере портала открытых данных правительства Москвы / В. И. Кузовлев, А. О. Орлов // Инженерный вестник. – 2014. – № 8. – С. 7–14.

13. Назаров, А. О. Выбор вида функции принадлежности в нечеткой модификации алгоритма COBWEB для задачи формирования пользовательских ролей / А. О. Назаров, И. В. Аникин // Вестник Казанского государственного технического университета им. А.Н. Туполева. – 2014. – № 2. – С. 214–219.

14. Пантелеев, А. В. Градиентные методы оптимизации в машинном обучении идентификации параметров динамических систем / А. В. Пантелеев, А. В. Лобанов // Моделирование и анализ данных. – 2019. – № 4. – С. 88–99.

15. Сафронов, В. В. Методы построения функций принадлежности / В. В. Сафронов, Ю. В. Ведерников // Информационные технологии. – 2007. – № 11. – С. 8–11.

References

1. Azhder, T. B., Gureeva, O. A. Sistemy podderzhki prinyatiya resheniy i informatsionnye sistemy [Decision support systems and information systems]. *Uralskiy nauchnyy vestnik* [Ural Scientific Bulletin], 2019, vol. 6, no. 3, pp. 46–48.

2. Barinov, A. I., Katasyova, D. V., Katasyov, A. S. Ispolzovanie modeli nechetkikh neyronnykh setey dlya formirovaniya bazy znaniy po opredeleniyu fishingovykh saytov [Using a fuzzy neural network model to form a knowledge base for identifying fishing sites]. *Vestnik Tekhnologicheskogo universiteta* [Bulletin of Technological University], 2020, vol. 23, no. 10, pp. 64–67.

3. Bashlykov, A. A. Rol cheloveka-operatora kak litsa, prinimayushchego resheniya, i elementa intellektualnoy sistemy upravleniya slozhnymi tekhnologicheskimi obektami [The role of a human operator as a decision maker and an element of an intelligent control system for complex technological objects]. *Avtomatizatsiya, telemekhanizatsiya i svyaz v neftyanoy promyshlennosti* [Automation, Telemechanization and Communication in the Oil Industry], 2016, vol. 12, pp. 10–18.

4. Glova, V. I., Katasyov, A. S., Kornilov, G. S. Prednastroyka i optimizatsiya parametrov nechetkoy neyronnoy seti pri formirovanii baz znaniy ekspertnykh sistem [Presetting and optimization of fuzzy neural network parameters when forming knowledge bases of expert systems]. *Informatsionnye tekhnologii* [Information Technologies], 2010, no. 5, pp. 15–19.

5. Gridina, N. V., Evdokimov, I. A., Solodovnikov, V. I. Postroenie gibridnykh neyronnykh setey s ispolzovaniem elementov nechetkoy logiki [Constructing hybrid neural networks using fuzzy logic elements]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial intelligence and decision making], 2019, no. 2, pp. 91–97.

6. Dzhumanov, O. I., Kholmonov, S. M. Metody optimizatsii identifikatsii nestatsionarnykh obektov na osnove neyro-nechetkoy seti s nastroykoy parametrov vychislitelnykh skhem [Methods for optimizing the identification of non-stationary objects based on a neuro-fuzzy network with setting the parameters of computational circuits]. *Problemy vychislitelnoy i prikladnoy matematiki* [Problems of Computational and Applied Mathematics], 2016, no. 3 (5), pp. 20–27.

7. Zvonkov, V. B., Popov, A. M. Sravnitelnoe issledovanie klassicheskikh metodov optimizatsii i geneticheskikh algoritmov [Comparative study of classical optimization methods and genetic algorithms]. *Vestnik Sibirskogo gosudarstvennogo aerokosmicheskogo universiteta imeni akademika M.F. Reshetneva* [Bulletin of the Siberian State Aerospace University named after Academician M.F. Reshetnev], 2013, no. 4 (50), pp. 23–27.

8. Klyueva, A. R. Issledovanie i otsenka sostoyaniya tekhnicheskikh obektov na osnove kompyuternogo modelirovaniya [Technical objects state research and assessment based on computer simulation]. *Inzhenernyy vestnik Dona* [Engineering Bulletin of the Don], 2018, no. 4 (51), pp. 120–132.

9. Komartsov, L. G., Kadnikov, D. S. Issledovanie geneticheskikh algoritmov dlya obucheniya mnogosloynogo perseptrona [Study of genetic algorithms for training multilayer perceptron]. *Neyrokompyutery: razrabotka, primenenie* [Neurocomputers: Development, Application], 2010, no. 12, pp. 12–19.

10. Koroteev, M. V. Formy funktsii prinadlezhnosti lingvisticheskikh peremennykh ekonomicheskikh pokazateley [Forms of the membership function of linguistic variables of economic indicators]. *Audit i finansovyy analiz* [Audit and Financial Analysis]. 2012, no. 2, pp. 239–244.

11. Kravchenko, T. K., Seredenko, N. N. Sozdanie sistem podderzhki prinyatiya resheniy: integratsiya preimushchestv otdelnykh podkhodov [Constructing decision support systems: integrating the benefits of separate approaches]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial intelligence and decision making], 2012, no. 1, pp. 39–47.

12. Kuzovlev, V. I., Orlov, A. O. Povyshenie kachestva dannykh s ispolzovaniem metodiki poiska anomaliy na primere portala otkrytykh dannykh pravitelstva Moskvy [Improving the quality of data using the anomaly search technique on the example of the open data portal of the Moscow government]. *Inzhenernyy vestnik* [Engineering Bulletin], 2014, no. 8, pp. 7–14.

13. Nazarov, A. O., Anikin, I. V. Vybor vida funktsii prinadlezhnosti v nechetkoy modifikatsii algoritma COBWEB dlya zadachi formirovaniya polzovatelskikh roley [Choosing the type of membership function in the fuzzy modification of the COBWEB algorithm for user roles generating problem]. *Vestnik Kazanskogo gosudarstvennogo tekhnicheskogo universiteta imeni A.N. Tupoleva* [Bulletin of the Kazan State Technical University named after A.N. Tupolev], 2014, no. 2, pp. 214–219.

14. Panteleev, A. V., Lobanov, A. V. Gradientnye metody optimizatsii v mashinnom obuchenii identifikatsii parametrov dinamicheskikh sistem [Gradient optimization methods in machine learning for identification of dynamic system parameter]. *Modelirovanie i analiz dannykh* [Modeling and Data Analysis], 2019, no. 4, pp. 88–99.

15. Safronov, V. V., Vedernikov, Yu. V. Metody postroeniya funktsiy prinadlezhnosti [Methods of membership functions constructing]. *Informatsionnye tekhnologii* [Information Technologies], 2007, no. 11, pp. 8–11.

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ МАШИН, КОМПЛЕКСОВ И КОМПЬЮТЕРНЫХ СЕТЕЙ

DOI 10.54398/2074-1707_2022_1_77
УДК 004:9

ЭЛЕМЕНТЫ НЕЧЕТКО-ГРАНУЛЯРНЫХ ВЫЧИСЛЕНИЙ В ПРИЛОЖЕНИИ К ЗАДАЧАМ АНАЛИЗА ТЕХНИЧЕСКОЙ НАДЕЖНОСТИ СИСТЕМ РАСПРЕДЕЛЕННОЙ ОБРАБОТКИ ДАННЫХ

Статья поступила в редакцию 15.01.2022, в окончательном варианте – 30.01.2022.

Михайличенко Антон Валерьевич, Военная академия связи им. Маршала Советского Союза С.М. Буденного, 194064, Российская Федерация, г. Санкт-Петербург, пр. Тихорецкий, 3,
адъюнкт, e-mail: antoxa9111603538@gmail.com

Паращук Игорь Борисович, Военная академия связи им. Маршала Советского Союза С.М. Буденного, 194064, Российская Федерация, г. Санкт-Петербург, пр. Тихорецкий, 3,
доктор технических наук, профессор, e-mail: shchuk@rambler.ru

Исследованы важные особенности применения нечетко-гранулярных вычислительных методов для анализа технической надежности систем распределенной обработки данных. Нечетко-гранулярные вычисления нацелены на уточнение (верификацию) неточных, «зашумленных» нечетких исходных данных большой размерности, необходимых для решения подобных задач анализа. Рассмотрены этапы, физическая сущность и математические аспекты подходов, реализующих информационное гранулирование и нечетко-гранулярные вычисления, позволяющих нивелировать нечеткость, «зашумление», неупорядоченность и неформализованность при формировании исходных данных для моделирования и текущей оценки параметров технической надежности систем такого класса. Применение предложенных нечетко-гранулярных методов в практике разработчиков и инженеров-исследователей позволит повысить объективность задания исходных данных как на этапе синтеза системы показателей технической надежности, так и на этапе формирования математической модели, что в итоге позволит повысить адекватность и точность анализа надежности реальных сложных информационных систем.

Ключевые слова: техническая надежность, нечетко-гранулярные вычисления, нечеткие исходные данные, метод, анализ, система распределенной обработки данных, показатель

ELEMENTS OF FUZZY-GRANULAR COMPUTING IN APPLICATION TO THE TASKS OF ANALYZING THE TECHNICAL RELIABILITY OF DISTRIBUTED DATA PROCESSING SYSTEMS

The article was received by editorial board 15.01.2022, in the final version – 30.01.2022

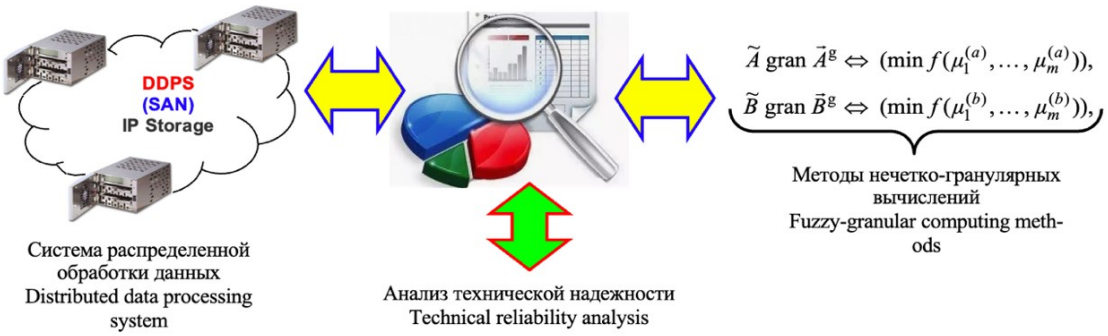
Mikhailichenko Anton V., Military Academy of Communications named after Marshal of the Soviet Union S.M. Budyonny, 3 Tikhoretsky Ave., St. Petersburg, 194064, Russian Federation,
post-graduate student, e-mail: antoxa9111603538@gmail.com

Parashchuk Igor B., Military Academy of Communications named after Marshal of the Soviet Union S.M. Budyonny, 3 Tikhoretsky Ave., St. Petersburg, 194064, Russian Federation,
Doct. Sci. (Engineering), Professor, e-mail: shchuk@rambler.ru

The important features of the application of fuzzy-granular computational methods in the interests of analyzing the technical reliability of distributed data processing systems are investigated. Fuzzy-granular computing are aimed at clarifying (verifying) inaccurate, “noisy” fuzzy source data of large dimension necessary for solving such analysis problems. The stages, physical essence and mathematical aspects of approaches implementing information granulation and fuzzy-granular calculations are considered, allowing to level out the fuzziness, “noise”, disorder and informality in the formation of initial data for modeling and current evaluation of the parameters of technical reliability of systems of this class. The application of the proposed fuzzy-granular methods in the practice of developers and research engineers will increase the objectivity of the initial data assignment both at the stage of synthesis of the system of technical reliability indicators and at the stage of formation of a mathematical model, which, ultimately, will increase the adequacy and accuracy of the reliability analysis of real complex information systems.

Keywords: technical reliability, fuzzy-granular computing, fuzzy source data, method, analysis, distributed data processing system, indicator

Graphical annotation (Графическая аннотация)



Введение. Неуклонное и всестороннее совершенствование технологий и средств автоматизации оперативной обработки, хранения и интеллектуального анализа больших массивов информации является нормой современного эволюционного развития IT-инфраструктуры страны и обязательной процедурой в рамках «цифровой трансформации» многих аспектов деятельности нашего государства.

В этой связи проблемы оперативной обработки и доведения до лиц, принимающих решения, достоверной информации во всех звеньях управления промышленностью, финансами, обороной, здравоохранением и другими сферами интересов страны приобретают особую актуальность. С решением именно этих проблем связана специфика перехода на качественно новый уровень управления, все более существенное место в целевых программах развития государства отводится завершению создания и совершенствованию алгоритмов функционирования единого информационного пространства Российской Федерации.

Одним из важнейших элементов единого информационного пространства, узловым компонентом IT-инфраструктуры, являются распределенные системы обработки данных (РСОД), к которым относятся как системы распределенных баз данных и центры обработки данных, так и несложные, с точки зрения топологии, при этом мощные вычислительные системы с распределенными ресурсами (многоядерные компьютерные системы), локальные ведомственные (и глобальная) сети, кластеры и «облачные» платформы [1–3].

Пользователь РСОД получает возможность работать с базами и хранилищами данных, прикладными процессами, программами и сервисами, расположенными в нескольких взаимосвязанных оконечных подсистемах или комплексах хранения и обработки данных. Примером РСОД, именуемых в зарубежной научной литературе DDPS (Distributed Data Processing Systems), могут служить как сети хранения данных – SAN (Storage Area Network), так и так называемые IP Storage-решения – комплексы хранения и обработки данных, накопители, доступ к которым осуществляется с использованием IP-технологий [4–6].

При этом распределенные вычислительные ресурсы РСОД должны обладать технологическими возможностями поиска и, главное, возможностями многомерного анализа данных, позволяющего математически корректно обрабатывать информацию и обеспечивать своевременность, достоверность и безопасность ее доведения до должностных лиц и иных пользователей [7–9].

Среди проблем и препятствий, стоящих на пути создания современных РСОД, среди трудностей, сопровождающих процесс их совершенствования, важное место, безусловно, принадлежит проблеме анализа и обеспечения технической надежности систем такого класса.

С учетом этой проблемы, при построении и совершенствовании многих сложных управляемых технических (информационных) систем, на наш взгляд, имеет право на жизнь гипотеза о рациональности применения новых методологических подходов, ориентированных на анализ технической надежности РСОД с учетом современных и перспективных математических методов вычислений, методов, способных обеспечить точность и оперативность анализа технической надежности РСОД даже в условиях неопределенности, нечеткости, неточности («зашумленности») исходных данных большой размерности. Одному из возможных подходов в рамках решения подобных задач посвящена настоящая статья.

Постановка задачи исследования. Распределенные системы обработки данных представляют собой развернутые на местности и взаимосвязанные комплексы программно-аппаратных средств для поиска, многопараметрической обработки и многомерного анализа информации. Они служат основой создания высокопроизводительной и отказоустойчивой инфраструктуры, отвечающей за оперативную обработку, хранение и интеллектуальный анализ больших объемов такой

информации. При этом данные системы нацелены на реализацию целого комплекса операций с информацией (попадающих под стандартное определение «обработка данных» или «обработка информации»), проводимых на взаимоувязанных, однако независимых вычислительных машинах.

Распределенные системы обработки данных процессуально построены на основе взаимосвязи функций отдельных комплексов хранения и обработки данных (КХОД). С точки зрения современных технологий вычислений и хранения данных, функции отдельных КХОД часто делят на «облачные», «туманные» и «граничные». С учетом этого можно охарактеризовать общую структуру РСОД как взаимоувязанную совокупность стационарного центрального КХОД, региональных и периферийных КХОД.

При этом центральный КХОД выполняет функции так называемых «облачных» вычислений (Cloud Computing). Стационарные региональные КХОД выполняют задачи, в мировой практике называемые «туманными» вычислениями (Fog Computing), при которых хранение и обработка информации осуществляется либо в рамках самой региональной КХОД, либо в региональной локальной сети, связывающей региональные КХОД и оконечные устройства абонентов. Типичный пример региональных КХОД, которые приближены к реальным пользователям, – средства и комплексы IoT (Internet of Things). Наконец, нижнее звено иерархии РСОД – современные мобильные, либо мини, модульные, «чемоданные» КХОД, ориентированные на «граничные», периферийные вычисления и хранение данных (Edge Computing) [10].

Важное достоинство таких компонентов РСОД – масштабируемость, а общая идея их создания – максимальное «приближение» вычислительных мощностей и ресурсов хранения данных к пользователю, в эталонном варианте – на точку сбора и «потребления» информации, рядом с оконечными устройствами пользователей РСОД.

Особенно актуально это для РСОД, строящихся в интересах обеспечения вычислительными мощностями и ресурсами хранения данных труднодоступных регионов (например, в Арктике), а также тропических и пустынных районов ведения, например, геологоразведки. Мобильные КХОД способны работать в широком диапазоне температур и даже в зонах с высокой вулканической и сейсмоактивностью. Кроме того, мобильные КХОД в большинстве – герметичны, а потому их можно перемещать, например, по грунтовым дорогам, морским (речным) или авиатранспортом.

Помимо удобства транспортировки, у мобильных КХОД, входящих в состав РСОД, есть еще ряд преимуществ:

- доступность вычислительных мощностей и ресурсов хранения данных для пользователей почти в любом месте их повседневной деятельности;
- модульная конструкция с точки зрения технологических решений.
- небольшие размеры и отсутствие требований к развитой инженерной инфраструктуре;
- невысокие энергетические затраты;
- низкая стоимость развертывания по сравнению с традиционными КХОД и дата-центрами;
- темпы изготовления и поставки (в разы быстрее строительства стандартных стационарных элементов РСОД).

Вместе с тем, наряду с безусловными достоинствами: прозрачностью, открытостью, переносимостью приложений, гибкостью, масштабируемостью и безопасностью, к ключевым характеристикам отдельных КХОД и РСОД в целом по праву относят их техническую надежность (ТН).

Под технической надежностью РСОД понимается свойство систем сохранять длительное время в заранее заданных пределах (диапазоне) значения всех своих параметров, характеризующих их способность выполнять требуемые функции в определенных режимах и условиях применения, технического обслуживания и транспортировки. Численно ТН чаще всего описывают средним временем наработки на отказ или вероятностью отказа в работе (вероятностью безотказной работы) и измеряют на определенном временном интервале [11–14].

Несмотря на большое количество исследований в этой области, пока не выработаны единые подходы к созданию достоверных, оперативных механизмов и математических методов анализа ТН сложных технических систем с распределенной структурой. Отсутствуют механизмы и методы, позволяющие максимально точно, в отведенные сроки и с учетом неопределенности наблюдаемых частных параметров ТН, осуществить анализ надежности отдельных КХОД и РСОД в целом. При этом важный объективный фактор – неопределенность, нечеткость исходных данных для анализа ТН РСОД, обусловливает не только теоретическую возможность, но и практическую необходимость, фактическую рациональность привлечения для задач анализа надежности РСОД новых методов. Методов, позволяющих нивелировать неопределенность, нечеткость исходных данных, например, нечетко-гранулярных методов (НГМ), иногда называемых алгоритмами fuzzy-granular computing – нечетко-гранулярных вычислений (НГВ) [15].

Методологические аспекты анализа технической надежности РСОД на основе методов нечетко-гранулярных вычислений. Методы нечетко-гранулярных вычислений потенциально способны решать задачи идентификации и анализа большого количества (больших массивов, объемов) нечетко заданных параметров ТН таких сложно структурированных, распределенных и динамических систем, как РСОД, так как могут работать с исходными данными, которые часто носят не только нечеткий, но и неформализованный, зашумленный, неупорядоченный характер.

Нечетко-гранулярные методы учитывают, что нечеткие исходные данные о значениях параметров ТН РСОД могут быть описаны (наблюдаемы) как точные, формализованные и однозначные лингвистические переменные и функции принадлежности, а могут быть избыточными и неточными, неупорядоченными и неформализованными. О таких данных в технической литературе часто говорят, как о «зашумленных» данных [16, 17].

Речь идет о том, что могут быть сформированы (наблюдаемы) массивы данных, математически, в терминах теории нечетких множеств, характеризующие степень принадлежности или непринадлежности конкретного параметра ТН РСОД к определенному конкретному множеству допустимых значений этого параметра. При этом, ввиду большого объема таких данных, существующих ошибок наблюдения и измерения (идентификации), неупорядоченности и слабой их формализованности, такие данные имеют «разброс» значений. Причем в этом случае трудно, а иногда и невозможно идентифицировать точные границы разрозненных множеств (массивов) данных. Именно об этих случаях идет речь, когда подразумевается, что конкретные данные не только нечеткие, но и неточны, «зашумлены».

Нечетко-гранулярные методы позволяют преодолеть проблемы анализа таких «зашумленных» данных, оценивания нечетко сформулированных (наблюдаемых) текущих и прогностических значений параметров ТН РСОД. Для этого в рамках НГМ предлагается использовать информационное гранулирование – слияние массивов нечетко заданных неточных, «зашумленных» исходных данных в группы (информационные гранулы) по принципу семантического и функционального сходства, а также применить математически корректные НГВ, нечетко-гранулярную обработку этих данных.

Предполагается, что НГМ для задачи оценивания параметров ТН РСОД могут быть использованы на двух фазах анализа, в рамках двух последовательных стадий НГВ:

Фаза информационного гранулирования – слияние больших массивов нечетко заданных неточных, «зашумленных» исходных данных о значениях параметров ТН РСОД в информационные гранулы по принципу функционального сходства. На этой фазе данные группируются в гранулы (множества) по принципу сходства функций принадлежности – минимального численного расстояния между значениями большого количества функций принадлежности, которые описывают конкретный нечетко заданный параметр ТН РСОД.

Иными словами, в нашем случае гранулой называется группа параметров ТН РСОД, с точки зрения математики представляющая собой динамическую целостную информационную структуру. Это множество информационных объектов, объединяемых неразличимостью, сходством, близостью (близостью значений функций принадлежности). При этом НГВ как совокупность методов математической обработки и преобразования информационных гранул могут быть использованы в качестве математического и методологического инструмента для обработки нечеткой информации [15, 18–20].

Примером может служить формат представления информационных гранул, индуцированных нечеткостью [18]. Нечеткая гранула (в нашем случае нечеткое множество параметров ТН РСОД) может быть представлена как произведение независимых скалярных экспоненциальных функций. При этом гранула рассматривается, в частности, как тензорное произведение векторов, представляющих собой элементы множества упорядоченных пар – нечетких множеств параметров ТН РСОД и их функций принадлежности, т.е.

$$\{\tilde{\gamma} \,\big|\, \mu^{(\gamma)}\} \to (\tilde{\gamma} \otimes \mu^{(\gamma)}), \tag{1}$$

где $\tilde{\gamma}$ – нечеткое множество параметров ТН РСОД; $\mu^{(\gamma)}$ – функция принадлежности нечеткого множества; $\otimes$ – символ тензорного произведения.

При этом нечеткое множество как информационная гранула – объект, элементы которого (α-уровни) связаны иерархической структурой, свойства которой определяются матрицей семантического и функционального сходства (сходства расстояний между значениями функций принадлежности, которые описывают конкретный нечетко заданный параметр ТН РСОД). Иными словами, определяется, к какому нечеткому множеству данный конкретный нечетко заданный параметр

ТН РСОД математически (функционально) «тяготеет». Подобные задачи ранее решались в рамках кластеризации [20, 21].

Группировка в гранулы (множества), например, в две гранулы $\vec{A}^{\mathrm{g}}$ и $\vec{B}^{\mathrm{g}}$, по принципу сходства функций принадлежности осуществляется на основе минимального численного расстояния между значениями большого количества M ($m = \overline{2, M}$; M может принимать значения от 2 до 100) функций принадлежности, которые описывают конкретный нечетко заданный параметр ТН РСОД

$$\widetilde{A}\ \mathrm{gran}\ \vec{A}^{\mathrm{g}} \Leftrightarrow\ (\min f(\mu_1^{(a)}, \ldots, \mu_m^{(a)})), \tag{2}$$

$$\widetilde{B}\ \mathrm{gran}\ \vec{B}^{\mathrm{g}} \Leftrightarrow\ (\min f(\mu_1^{(b)}, \ldots, \mu_m^{(b)})), \tag{3}$$

где a и b – элементы нечетких множеств $\widetilde{A}$ и $\widetilde{B}$ соответственно; $\mu_m^{(a)}$ и $\mu_m^{(b)}$ – функции принадлежности конкретных элементов нечетких множеств $\widetilde{A}$ и $\widetilde{B}$ соответственно; gran – символ информационного гранулирования, физический и математический смысл которого заключается в группировании данных на основе сходства функций принадлежности – минимального численного (функционального f) расстояния $(\min f(\mu_1^{(a)}, \ldots, \mu_m^{(a)}))$ и $(\min f(\mu_1^{(b)}, \ldots, \mu_m^{(b)}))$ между значениями большого количества m функций принадлежности, которые описывают нечетко заданные параметры (группы параметров) ТН РСОД [20, 21].

В итоге получаем, что нечеткие множества $\widetilde{A}$ и $\widetilde{B}$, характеризующие состав групп нечетко заданных параметров ТН РСОД, соответствуют гранулам $\vec{A}^{\mathrm{g}}$ и $\vec{B}^{\mathrm{g}}$, которые можно записать в виде

$$\vec{A}^{\mathrm{g}} = (a_1\, \mu_1^{(a)}; \ldots; a_m\, \mu_m^{(a)}), \tag{4}$$

$$\vec{B}^{\mathrm{g}} = (b_1\, \mu_1^{(b)}; \ldots; b_m\, \mu_m^{(b)})\,. \tag{5}$$

Фаза собственно НГВ – математическая обработка информационных гранул с целью преобразования характеризующих их неточных, «зашумленных», неупорядоченных и неформализованных нечетких исходных данных большой размерности (избыточных данных) к виду, пригодному для осуществления достоверной параметрической оценки ТН РСОД [20, 21].

На этой фазе производится операция гранулярного суммирования над векторами (гранулами) $\vec{A}^{\mathrm{g}}$ и $\vec{B}^{\mathrm{g}}$ (иногда записывают $\vec{A}^{\mathrm{g}(a)}$ и $\vec{B}^{\mathrm{g}(b)}$) и вычисление функции следа гранулярной суммы.

Операция гранулярного суммирования – это, в сущности, операция нечетко-гранулярного математического объединения неточно заданных, «зашумленных» нечетких множеств – исходных данных, характеризующих, в нашем случае, степень уверенности экспертов в количественных значениях параметров ТН РСОД

$$\vec{C}^{\mathrm{g}(a,b)} = \vec{A}^{\mathrm{g}(a)} +_{\mathrm{g}} \vec{B}^{\mathrm{g}(b)}, \tag{6}$$

где $+_{\mathrm{g}}$ – арифметическая операция гранулярного суммирования [15, 20, 21].

Гранулярная сумма подразумевает дефаззификацию (преобразование нечеткого множества в четкое число) элементов гранул нечетких множеств и вычисление функции следа гранулярной суммы [19–21].

Например, для гранулы $\vec{A}^{\mathrm{g}(a)} = (a_1\, \mu_1^{(a)}; a_2\, \mu_2^{(a)}; a_3\, \mu_3^{(a)})$, содержащей три элемента – три значения нечетко заданного параметра ТН РСОД, входящего в группу параметров ТН (множество) $\vec{A}^{\mathrm{g}(a)}$, выражение для дефаззификации будет иметь вид

$$a^{(\mathrm{def})} = \left(\sum_{i=1}^{3} a_i\, \mu_i^{(a)}\right) \Big/ \left(\sum_{i=1}^{3} \mu_i^{(a)}\right). \tag{7}$$

Затем осуществляется минимизация функции следа гранулярной суммы как итог реализации преобразования нечетких и «зашумленных» переменных к точному количественному виду.

Например, для нашего примера, для гранулы $\vec{A}^{\mathrm{g}(a)}$ как группы параметров ТН, содержащей три элемента – три значения нечетко заданного параметра ТН РСОД, функцию следа гранулярной суммы находят с помощью выражения

$$\text{tr } a = \left(\sum_{i=1}^{3} a_i^2 + \mu_i^2 \right)^{\frac{1}{2}}. \quad (8)$$

Выполняется операция над векторами (гранулами) $\vec{A}^{g(\mu)}$ и $\vec{B}^{g(\mu)}$. Это операция минимизации – min ($\vec{A}^{g(\mu)}$, $\vec{B}^{g(\mu)}$), т.е.

$$\vec{C}^{g(\mu)} = \min (\vec{A}^{g(\mu)}, \vec{B}^{g(\mu)}). \quad (9)$$

Итоговый результат НГВ находят из выражения [19–21]

$$\vec{C}^{g} = [\vec{C}^{g(a,b)} \; \vec{C}^{g(\mu)}] = \begin{pmatrix} a_1 +_g b_1 & \min(\mu_1^{(a)}, \mu_1^{(b)}) \\ & \vdots \\ a_m +_g b_m & \min(\mu_m^{(a)}, \mu_m^{(b)}) \end{pmatrix}. \quad (10)$$

Физический смысл результата использования НГМ для задачи анализа ТН РСОД заключается в получении верифицированного (точного, «незашумленного») значения конкретного элемента исходного нечеткого множества, содержащего информацию о количественном значении соответствующего конкретного оцениваемого параметра ТН систем такого класса.

При этом в рамках решения задачи анализа ТН РСОД применение НГМ может быть успешно реализовано на нескольких этапах анализа: на этапе синтеза безызбыточной системы показателей ТН РСОД, на этапе получения экспертных количественных оценок значений конкретного параметра ТН РСОД, а также в рамках синтеза элементов матрицы переходных вероятностей – ключевого компонента марковской математической модели процесса смены состояний показателей ТН в динамике функционирования РСОД [22].

Заключение. Таким образом, рассмотренные основы применения нечетко-гранулярных методов позволяют повысить достоверность анализа технической надежности распределенных систем обработки информации за счет уточнения (верификации) неточных, «зашумленных» нечетких исходных данных большой размерности. Эти методы, основанные на информационном гранулировании и нечетко-гранулярных вычислениях, позволяют нивелировать нечеткость, «зашумление», неупорядоченность и неформализованность при формировании исходных данных для моделирования и текущей оценки параметров технической надежности систем такого класса. Это, в свою очередь, позволяет повысить объективность задания этих исходных данных как на этапе синтеза системы показателей ТН, на этапе формирования математической модели, а также на этапе оценивания этих показателей, что в итоге служит повышению адекватности и точности анализа надежности реальных сложных информационных систем.

Библиографический список

1. Влацкая, И. В. Распределенная обработка информации : учебное пособие / И. В. Влацкая, С. И. Сормов. – Оренбур г: ИПК ГОУ ОГУ, 2010. – 146 с.

2. Бабичев, С. Л. Распределенные системы : учебное пособие для вузов / С. Л. Бабичев, К. А. Коньков. – Москва : Юрайт, 2019. – 507 с.

3. Маглинец, Ю. А. Анализ требований к автоматизированным информационным системам / Ю. А. Маглинец. – Москва : Бином. Лаборатория знаний: Интернет-Университет информационных технологий, 2008. – 199 с.

4. Van Steen, M. Distributed Systems / M. Van Steen, A. S. Tanenbaum. – 3rd ed. – Distributed-Systems.Net, 2017. – Режим доступа: https://www.distributed–systems.net/index.php/books/ds3/, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 13.01.2022).

5. Гусейнов, А. А. Исследование распределенной обработки данных на примере системы Hadoop / А. А. Гусейнов, И. А. Бочкова // Актуальные проблемы авиации и космонавтики. – 2016. – № 1. – С. 606–608.

6. Puder, A. Distributed systems architecture: a middleware approach / A. Puder, K. Römer, F. Pilhofer. – San Francisco : Morgan Kaufmann Publishers is an imprint of Elsevier, 2006. – 342 p.

7. Андреев, Д. В. Универсальные логические модули для обработки многозначных и континуальных данных / Д. В. Андреев. – Ульяновск : УлГТУ, 2010. – 234 с.

8. Большаков, А .А. Методы обработки многомерных данных и временных рядов : учебное пособие / А. А. Большаков, Р. Н. Каримов. – Москва : Горячая линия-Телеком, 2007. – 522 с.

9. Сазонов, В. В. Математическое обеспечение АСУ войсками : учебное пособие / В. В. Сазонов, И. Б. Паращук, В. А. Логинов, В. В. Елизаров ; под ред. проф. И. Б. Паращука. – Санкт-Петербург : ВАС, 2018. – 256 с.

10. Дмитриев, К. А. Discover the Edge: современные решения для задач будущего / К. А. Дмитриев // ИнформКурьер-Связь. – 2019. – № 3. – С. 58–59.

11. Громов, Ю. Ю. Надежность информационных систем : учебное пособие / Ю. Ю. Громов, О. Г. Иванова, Н. Г. Мосягина, К. А. Набатов. – Тамбов : ГОУ ВПО ТГТУ, 2010. – 160 с.

12. Крюкова, Е.С. Вопросы оценки надежности современных систем хранения данных для мобильных дата-центров / Е. С. Крюкова, В. В. Ткаченко, А. В. Михайличенко, И. Б. Паращук // Наукоемкие технологии в космических исследованиях Земли. – 2021. – Т. 13, № 5. – С. 86–95.

13. Гуров, С. В. Надежность систем при неполной информации / С. В. Гуров, Л. В. Уткин. – Санкт-Петербург : Любавич, 1999. – 160 с.

14. Балакирев, В. С. Надежность и диагностика автоматизированных систем : учебное пособие для вузов / В. С. Балакирев, А. А. Большаков. – Санкт-Петербург : Изд-во Политехн. ун-та, 2018. – 144 с.

15. Бутакова, М. А. Элементы теории гранулярных вычислений с нечеткими приближенными информационными гранулами / М. А. Бутакова, А. Н. Гуда, О. В. Иванченко, Е. В. Карпенко // Вестник Ростовского государственного университета путей сообщения. – 2015. – № 4 (60). – С. 27–33.

16. Стаут, М. Обеспечение целостности данных в зашумленных средах / М. Стаут // Электронные компоненты. – 2012. – № 3. – С. 109–110.

17. Паращук, И. Б. Анализ зашумленных и неоднородных данных о значениях параметров надежности дата-центров / И. Б. Паращук, А. В. Михайличенко, Е. С. Крюкова // Современные технологии: актуальные вопросы теории и практики : сборник статей Международной научно–практической конференции. – Пенза : МЦНС «Наука и Просвещение», 2021. – С. 74–77.

18. Bargiela, A. Granular Computing: An Introduction / A. Bargiela, W. Pedrycz. – N. Y. : Springer, 2003. – 452 p.

19. Pedrycz, W. Handbook of Granular Computing / W. Pedrycz, A. Skowron, V. Kreinovich. – N.Y. : John Wiley & Sons, Ltd, 2008. – 283 p.

20. Минаев, Ю. Н. Гранулярный компьютинг в системе нечетких множеств на уровне тензорных гранул / Ю. Н. Минаев, О. Ю. Филимонова, Ю. И. Минаева // Проблемы информатизации и управления. – 2012. – № 4 (40). – С. 51–61.

21. Паращук, И. Б. Нейро-нечеткие сети и алгоритмы гранулярных вычислений в задачах интеллектуальной обработки данных для оценки надежности мобильных дата-центров / И. Б. Паращук, Н. В. Михайличенко, А. В. Михайличенко // Применение искусственного интеллекта в информационно-телекоммуникационных системах : сборник материалов научно-практической конференции (31 марта 2021 г., Военная академия связи, Санкт-Петербург). – Санкт-Петербург : ВАС, 2021. – С. 110–115.

22. Крюкова, Е.С. Математическая модель, предназначенная для оценки качества электронной библиотеки: синтез числа градаций пространства состояний / Е. С. Крюкова, И. Б. Паращук // Прикаспийский журнал: управление и высокие технологии. – 2020. – № 1 (49). – С. 121–131. – DOI: 10.21672/2074-1707.2020.49.4.121-131.

References

1. Vlatskaya I. V., Sormov S. I. *Raspredelennaya obrabotka informatsii : uchebnoe posobie* [Distributed information processing : a textbook]. Orenburg, IPK GOU OGU, 2010. 146 p.

2. Babichev S. L., Konkov K. A. *Raspredelennye sistemy : uchebnoe posobie dlya vuzov* [Distributed systems : a textbook for universities]. Moscow, Yurayt Publ., 2019. 507 p.

3. Maglinets Yu. A. *Analiz trebovaniy k avtomatizirovannym informatsionnym sistemam* [Analysis of requirements for automated information systems]. Moscow, Binom. Laboratoriya znaniy: Internet-Universitet Informatsionnykh tekhnologiy, 2008. 199 p.

4. Van Steen M., Tanenbaum A. S. *Distributed Systems.* 3rd ed. Distributed-Systems.Net, 2017. Available at: https://www.distributed-systems.net/index.php/books/ds3 (accessed 13.01.2022).

5. Guseynov A. A., Bochkova I. A. Issledovanie raspredelennoy obrabotki dannykh na primere sistemy Hadoop [The study of distributed data processing on the example of the Hadoop system]. *Aktualnye problemy aviatsii i kosmonavtiki* [Actual problems of aviation and cosmonautics], 2016, no. 1, pp. 606–608.

6. Puder, A. Römer K., Pilhofer F. *Distributed systems architecture: a middleware approach*. San Francisco, Morgan Kaufmann Publishers is an imprint of Elsevier, 2006. 342 p.

7. Andreev, D. V. *Universalnye logicheskie moduli dlya obrabotki mnogoznachnykh i kontinualnykh dannykh* [Universal logic modules for processing multi-valued and continuous data]. Ulyanovsk, Ulyanovsk State Technical University, 2010. 234 p.

8. Bolshakov, A. A., Karimov R. N. *Metody obrabotki mnogomernykh dannykh i vremennykh ryadov : uchebnoe posobie* [Methods of processing multidimensional data and time series : textbook]. Moscow, Goryachaya liniya-Telekom Publ., 2007. 522 p.

9. Sazonov, V. V., Parashchuk I. B. (ed.), Loginov V. A., Elizarov V. V. *Matematicheskoe obespechenie ASU voyskami : uchebnoe posobie* [Mathematical support of automated control systems by troops : a textbook]. Saint Petersburg, Military Academy of the Sygnal Corps, 2018. 256 p.

10. Dmitriev, K. A. Discover the Edge: sovremennye resheniya dlya zadach budushhego [Discover the Edge: modern solutions for the challenges of the future]. *InformKurer-Svyaz* [InformCurrier-Communication], 2019, no. 3, pp. 58–59.

11. Gromov, Yu. Yu., Ivanova, O. G., Mosyagina, N. G., Nabatov, K. A. *Nadezhnost informatsionnykh system : uchebnoe posobie* [Reliability of information systems : a textbook]. Tambov, 2010. 160 p.

12. Kryukova, E. S., Tkachenko, V. V., Mikhaylichenko, A. V., Parashchuk, I. B. *Voprosy otsenki nadezhnosti sovremennykh sistem khraneniya dannykh dlya mobilnykh data-tsentrov* [Issues of reliability assessment of modern

data storage systems for mobile data centers]. *Naukoemkie tekhnologii v kosmicheskikh issledovaniyakh Zemli* [High-tech technologies in space exploration of the Earth], 2021, vol. 13, no. 5, pp. 86–95.

13. Gurov, S. V., Utkin, L. V. *Nadezhnost sistem pri nepolnoy informatsii* [Reliability of systems with incomplete information]. Saint Petersburg, Lyubavich Publ., 1999. 160 p.

14. Balakirev, V. S., Bolshakov, A. A. *Nadezhnost i diagnostika avtomatizirovannykh system : uchebnoe posobie dlya vuzov* [Reliability and diagnostics of automated systems: a textbook for universities]. Saint Petersburg, Publishing House of Polytechnical University, 2018. 144 p.

15. Butakova, M. A., Guda, A. N., Ivanchenko, O. V., Karpenko, E. V. Elementy teorii granulyarnykh vychisleniy s nechetkimi priblizhennymi informatsionnymi granulami [Elements of the theory of granular computing with fuzzy approximate information granules]. *Vestnik Rostovskogo gosudarstvennogo universiteta putey soobshcheniya* [Bulletin of the Rostov State University of Railways], 2015, no. 4 (60), pp. 27–33.

16. Staut, M. Obespechenie tselostnosti dannykh v zashumlennykh sredakh [Ensuring data integrity in noisy environments]. *Elektronnye komponenty* [Electronic Components], 2012, no. 3, pp. 109–110.

17. Parashchuk, I. B., Mikhaylichenko, A. V., Kryukova, E. S. Analiz zashumlennykh i neodnorodnykh dannykh o znacheniyakh parametrov nadezhnosti data-tsentrov [Analysis of noisy and heterogeneous data on the values of data center reliability parameters]. *Sovremennye tekhnologii: aktualnye voprosy teorii i praktiki : sbornik statey Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Modern technologies: topical issues of theory and practice : collection of articles of the International Scientific and Practical Conference]. Penza, MCzNS «Nauka i Prosveshhenie», 2021, pp. 74–77.

18. Bargiela, A., Pedrycz, W. *Granular Computing: An Introduction*. N.Y., Springer, 2003. 452 p.

19. Pedrycz, W., Skowron, A., Kreinovich, V. *Handbook of Granular Computing*. N.Y., John Wiley & Sons, Ltd, 2008. 283 p.

20. Minaev, Yu. N., Filimonova, O. Yu., Minaeva, Yu. I. Granulyarnyy kompyuting v sisteme nechetkikh mnozhestv na urovne tenzornykh granul [Granular computing in a system of fuzzy sets at the level of tensor granules]. *Problemy informatizatsii i upravleniya* [Problems of informatization and management], 2012, no. 4 (40), pp. 51–61.

21. Parashchuk, I. B., Mikhaylichenko N. V., Mikhaylichenko A. V. Neyro-nechetkie seti i algoritmy granulyarnykh vychisleniy v zadachakh intellektualnoy obrabotki dannykh dlya otsenki nadezhnosti mobilnykh data-tsentrov [Neuro-fuzzy networks and granular computing algorithms in intelligent data processing tasks for assessing the reliability of mobile data centers]. *Primenenie iskusstvennogo intellekta v informatsionno-telekommunikatsionnykh sistemakh : sbornik materialov nauchno-prakticheskoy konferentsii (31 marta 2021 g., Voennaya akademiya svyazi, Sankt-Peterburg)* [Application of artificial intelligence in information and telecommunication systems : collection of materials of the scientific and practical conference (March 31, 2021, Military Academy of Communications, Saint Petersburg)]. Saint Petersburg, Military Academy of the Sygnal Corps, 2021, pp. 110–115.

22. Kryukova, E. S., Parashchuk, I. B. Matematicheskaya model, prednaznachennaya dlya otsenki kachestva elektronnoy biblioteki: sintez chisla gradatsiy prostranstva sostoyaniy [A mathematical model designed to assess the quality of an electronic library: synthesis of the number of gradations of the state space]. *Prikaspiyskij zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2020, no. 1 (49), pp. 121–131. DOI: 10.21672/2074-1707.2020.49.4.121-131.

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И ЗАЩИТА ИНФОРМАЦИИ

DOI 10.54398/2074-1707_2022_1_85
УДК 004.052

ПОВЫШЕНИЕ СКОРОСТИ ОБНАРУЖЕНИЯ ОШИБОК ПРИ ФОРМИРОВАНИИ ЦЕПОЧЕК БЛОКОВ ДАННЫХ НА ОСНОВЕ АНАЛИЗА ЧИСЛА СОВПАДЕНИЙ ХЕШЕЙ [1]

Статья поступила в редакцию 29.11.2021, в окончательном варианте – 31.01.2022.

Таныгин Максим Олегович, Юго-Западный государственный университет, 305004, Российская Федерация, г. Курск, ул. Челюскинцев, 19, корпус Б,
кандидат технических наук, доцент, заведующий кафедрой информационной безопасности, ORCID: 0000-0002-4099-1414, e-mail: tanygin@yandex.ru
Кулешова Елена Александровна, Юго-Западный государственный университет, 305004, Российская Федерация, г. Курск, ул. Челюскинцев, 19, корпус Б,
преподаватель кафедры программной инженерии, ORCID: 0000-0002-8270-564X, e-mail: lena.kuleshova.94@mail.ru
Митрофанов Алексей Васильевич, Юго-Западный государственный университет, 305004, Российская Федерация, г. Курск, ул. Челюскинцев, 19, корпус Б,
аспирант кафедры информационной безопасности, ORCID: 0000-0001-7200-6418, e-mail: mitro3000@rambler.ru
Гладилина Елена Юрьевна, Юго-Западный государственный университет, 305004, Российская Федерация, г. Курск, ул. Челюскинцев, 19, корпус Б,
студент, e-mail: elena.gladilina@inbox.ru

В статье исследуется практическая реализация системы контроля целостности и аутентичности информации на основе CBC-кодов с целью повышения скорости обнаружения ошибок при формировании цепочек блоков данных. В рамках реализации схемы проверки блоков данных предлагается формировать древовидные структуры информационных блоков путём анализа атрибутов последних, при этом факт возникновения ошибки определяется на основе числа и длины ветвей в такой древовидной структуре. В статье предложен метод определения источника сообщений, который основан на анализе содержимого имитовставки и индекса сообщения в последовательности между двумя служебными сообщениями: стартовым, являющимся корнем древовидной структуры, и стоповым, которое должно быть последним сообщением в ветви. В основе исследования лежит математическая модель формирования древовидной структуры, на основании которой получены рекуррентные зависимости для вероятности формирования ветвей от корневого сообщения и вероятности формирования различных значений имитовставок сообщений этих ветвей. В работе показано, что существует взаимосвязь между числом ветвей древовидной структуры, порождаемых совпадениями значений атрибутов информационных блоков, и вероятностью ошибки при определении источника. На основе полученных оценок в работе сформулировано правило обработки ветвей древовидной структуры для сообщений ограниченной длины. Экспериментальные исследования показали, что применение данного правила позволит увеличить долю полезной информации, обрабатываемой приёмником, на 2–5 % за счёт снижения числа переспрашиваемых в результате обнаруженных ошибок аутентификации блоков данных.

Ключевые слова: передача данных, приёмник сообщений, система аутентификации, цепочки блоков данных, скорость обнаружения ошибок

INCREASING THE SPEED OF ERROR DETECTION WHEN FORMING CHAINS OF DATA BLOCKS BASED ON THE ANALYSIS OF THE NUMBER OF HASH MATCHES

The article was received by the editorial board on 29.11.2021, in the final version – 31.01.2022.

Tanygin Maxim O., Southwest State University, building B, 19 Chelyuskintsev St., Kursk, 305004, Russian Federation,
Cand. Sci. (Engineering), Associate Professor, Head of the Department of Information Security, ORCID: 0000-0002-4099-1414, e-mail: tanygin@yandex.ru
Kuleshova Elena A., Southwest State University, building B, 19 Chelyuskintsev St., Kursk, 305004, Russian Federation,
Lecturer at the Department of Software Engineering, ORCID: 0000-0002-8270-564X, e-mail: lena.kuleshova.94@mail.ru

[1] Исследование выполнено в рамках реализации внутриуниверситетского гранта по программе развития ЮЗГУ (ПРИОРИТЕТ-2030) № ПР2030/2021-27.

Mitrofanov Alexey V., Southwest State University, building B, 19 Chelyuskintsev St., Kursk, 305004, Russian Federation,

Postgr. St. of the Department of Information Security, ORCID: 0000-0001-7200-6418, e-mail: mitro3000@rambler.ru

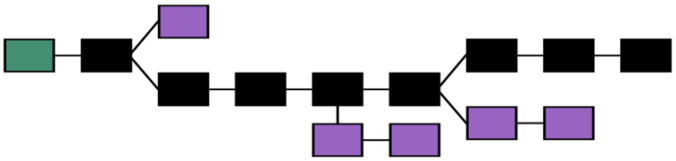
Gladilina Elena Yu., Southwest State University, building B, 19 Chelyuskintsev St., Kursk, 305004, Russian Federation,

student, e-mail: elena.gladilina@inbox.ru

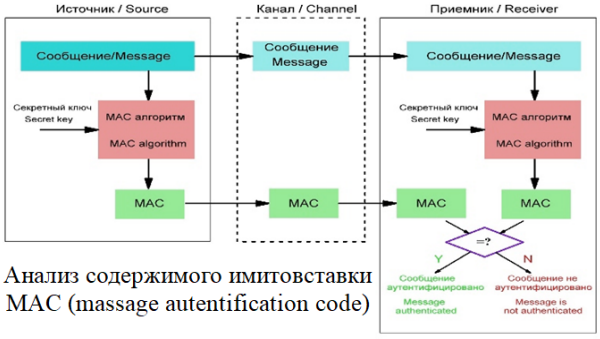
The article investigates the practical implementation of a system for monitoring the integrity and authenticity of information based on CBC codes in order to increase the speed of error detection when forming chains of data blocks. As part of the implementation of the scheme for checking data blocks, it is proposed to form tree-like structures of information blocks by analyzing the attributes of the latter, while the fact of an error is determined on the basis of the number and length of branches in such a tree-like structure. The article proposes a method for determining the source of messages, which is based on the analysis of the MAC content and the message index in the sequence between two service messages: start, which is the root of the tree structure, and stop, which should be the last message in the branch. The study is based on a mathematical model of the formation of a tree structure, on the basis of which recurrent dependencies for the probability of the formation of branches from the root message and the probability of the formation of different values of MAC messages of these branches are obtained. The paper shows that there is a relationship between the number of branches of the tree structure, generated by coincidences of the values of the attributes of information blocks, and the probability of error in determining the source. Based on the estimates obtained, a rule for processing branches of a tree structure for messages of limited length is formulated in the work. Experimental studies have shown that the application of this rule will increase the share of useful information processed by the receiver by 2–5 % by reducing the number of data blocks requested as a result of detected authentication errors.

Keywords: data transmission, message receiver, authentication system, data block chains, error detection rate

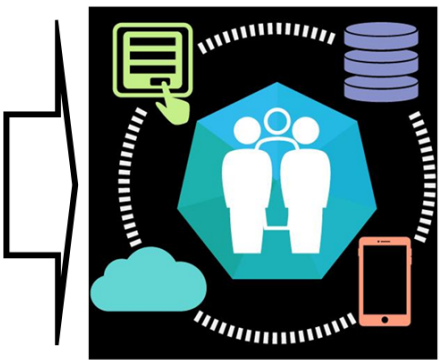
Graphical annotation (Графическая аннотация)



Формирование древовидной структуры информационных блоков
Formation of a tree structure of information blocks



Анализ содержимого имитовставки
MAC (massage autentification code)



Система контроля целостности
и аутентичности
Integrity and Authenticity
Control System

Введение. Широкое распространение автоматических систем управления технологическими процессами [1–3], формирование сетей, объединяющих в себя различные устройства мониторинга и управления техническими объектами [4, 5], влечёт за собой необходимость разработки специализированных протоколов передачи и обработки команд и данных. Особенности форматов данных и процесса их передачи, такие как временное разделение канала связи, передача данных в режиме широковещательных запросов [6, 7], вкупе с возросшими требованиями по обеспечению информационной безопасности таких систем [8, 9] обусловливает необходимость использования новых подходов к обеспечению аутентичности и целостности передаваемых и обрабатываемых данных. Использование ставших традиционными алгоритмов шифрования и помехоустойчивого кодирования не позволяет достичь требуемых показателей достоверности [10–13].

Кодирование передаваемых данных в режиме сцепления блоков является единственным средством повышения вероятности правильной аутентификации источника сообщений в условиях ограниченного размера поля имитовставки (англ. MAC – massage autentification code) [14, 15]. В то же время практическое применение такого подхода для выделения из общего потока команд

и данных сообщений от целевого источника порождает проблему обработки сложных структур [16]. Причиной этого является то, что при формировании структурированных последовательностей, в которых позиция каждого сообщения задана не напрямую индексом сообщения, а взаимным расположением относительно других сообщений (что позволяет обеспечить требуемую достоверность), в результате совпадения значений имитовставок различных сообщений вместо цепочки формируется сложная древовидная структура.

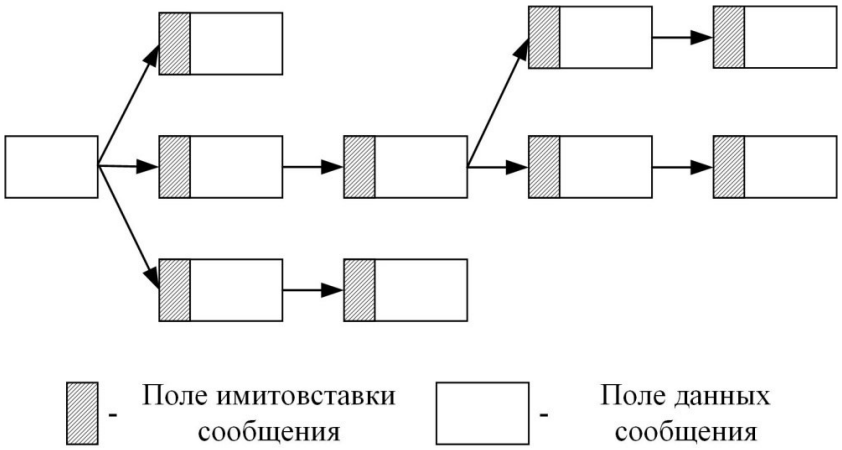


Рисунок 1 – Структура, формируемая в приемнике из-за совпадения значений имитовставок

В работах [17, 18] представлены оценки вероятности для числа и длины L ветвей в таких древовидных структурах, а также предложена математическая модель оценки достоверности процедуры аутентификации и её вычислительной сложности. Кроме того, различные варианты алгоритмов формирования древовидных структур показали, что количество операций сравнения имитовставок сообщений зависит от количества сформированных к текущему числу во внутренней памяти приёмника ветвей древовидной структуры [18]. Число таких сравнений непосредственно увеличивает вероятность возникновения коллизии, т. е. ситуации, при которой две или более ветви структуры удовлетворяют всем требованиям, предъявляемым к последовательности сообщений, чтобы быть ассоциированной с конкретным источником. Таким образом, число и размер ветвей в древовидной структуре оказываются связаны с достоверностью процедуры аутентификации источника, так как обе эти величины, как показано в [19], являются функцией вероятности случайного совпадения значений имитовставок, которая, в свою очередь, зависит от размера поля имитовставки в отдельном сообщении. Эта взаимосвязь является теоретической предпосылкой для обнаружения ошибок аутентификации не в результате анализа содержимого сообщений, входящих в ветви древовидной структуру, а на основе анализа числа таких цепочек на соответствующих этапах алгоритма.

Материалы и методы. Объектом исследования является метод определения источника сообщений, который основан на анализе содержимого имитовставки и индекса сообщения в последовательности между двумя служебными сообщениями: стартовым, являющимся корнем древовидной структуры, и стоповым, которое должно быть последним сообщением в ветви. В основе настоящего исследования лежит математическая модель формирования древовидной структуры, описанная в [19]. В ней, рассмотрев процесс формирования древовидной структуры, начиная с корня, получены рекуррентные зависимости для вероятности $p_r(k_r)$ формирования ровно k_r ветвей в позиции r от корневого сообщения и вероятности $p_r^{\text{h}}\left(k_r^{\text{h}}\right)$ формирования k_r^{h} различных значений имитовставок сообщений этих ветвей [16], которые вычисляются по (1):

$$p_r\left(k_r\right)=\sum_{l=k_r}^{|U|-n}\left[\frac{\left((U-n)/n\right)^l\times e^{-\frac{(U-n)}{n}}}{l!}\times\sum_{k_{r-1}=1}^{|U|-n}p_{r-1}^{\text{h}}\left(k_{r-1}^{\text{h}}\right)\frac{\left(k_{r-1}^{\text{h}}l2^{-H}\right)^l\times e^{-k_{r-1}^{\text{h}}l2^{-H}}}{l!}\right],$$

$$p_r^{\text{h}}\left(k_r^{\text{h}}\right)=\sum_{l=k_r^{\text{h}}}^{|U|-n}\left[p_r\left(l\right)\times\left(\left(2^{-H}\right)^{j-k_r^{\text{h}}}\prod_{k=1}^{k_r^{\text{h}}}\left(1-(k-1)2^{-H}\right)\right)\right]. \quad (1)$$

где U – общее число сообщений, обрабатываемых приёмником; H – размер поля имитовставки; k_r – число посторонних ветвей на расстоянии r от корня, n – длина цепочки сообщений.

В рассматриваемом подходе ошибка аутентификации источника блоков данных возникает, когда имитовставка, сформированная из какого-либо блока посторонней ветви древовидной структуры, совпадёт с имитовставкой соответствующего сообщения целевого источника. Вероятность этого определится по (2):

$$p_{\text{col}}(r)=\sum_{l=1}^{|U|-n}\left[p_r(l)\left(1-\left(1-2^{-H}\right)^l\right)\right], \tag{2}$$

где r – длина постороннней ветви.

Соответственно, функция зависимости вероятности возникновения ошибки от длины посторонней ветви, отсчитываемой от момента её отхода от основной ветви из сообщений целевого источника, рассчитывается по (3):

$$p\left(q_j\right)=C_{i-j}^{q_j}\left(p_{\text{col}}(j)\right)^{q_j}\left(1-p_{\text{col}}(j)\right)^{i-j-q_j},\ q_j=\overline{0\ldots} \tag{3}$$

где q_j – число возникших ошибок совпадения имитовставок; j – число сообщений в посторонней ветви, после которого возникла ошибка; i – позиция блока, от которого отходит посторонняя ветвь, в цепочке из n сообщений целевого источника.

Тогда вероятность возникновения ровно q ошибок совпадения значений имитовставок сообщений будет вычисляться по (4):

$$p^{\text{col}}(q)=\sum_{l=1}^{R[q]}\left(\prod_{k=1}^{l}p(q_k)\right),\ \sum_{k=1}^{l}q_k=q, \tag{4}$$

где $R[q]$ – число разбиений числа q на слагаемые q_k.

На основании рекуррентных формул (1) вероятность формирования в древовидной структуре q побочных ветвей, которые не привели к возникновению ошибки, определится как сумма произведений вероятностей по (5):

$$p^{\text{thr}}(q)=\sum_{l=1}^{R[q]}\left(\prod_{k=1}^{i-1}p_k(q_k)\right),\ \sum_{k=1}^{i-1}q_k=q. \tag{5}$$

Априорная вероятность формирования на определённом отдалении от корневого сообщения ровно q ветвей есть вероятность одновременного возникновения q' ошибок и q'' ветвей, ошибку не вызвавших. Исходя из того, что ошибка порождает дополнение посторонней ветви сообщениями целевого источника и удваивает число ветвей дерева, для q верно соотношение (6):

$$q=(q')^2+q''. \tag{6}$$

Априорная вероятность формирования двух ветвей в древовидной структуре вычисляется по (7):

$$p(2)=p^{\text{col}}(1)\cdot p^{\text{thr}}(0)+p^{\text{thr}}(2)p^{\text{col}}(0). \tag{7}$$

Для большего числа ветвей по (8):

$$\begin{aligned}
p(3)&=p^{\text{col}}(1)\cdot p^{\text{thr}}(1)+p^{\text{col}}(0)\cdot p^{\text{thr}}(3),\\
p(4)&=p^{\text{col}}(2)\cdot p^{\text{thr}}(0)+p^{\text{col}}(1)\cdot p^{\text{thr}}(2)+p^{\text{col}}(0)\cdot p^{\text{thr}}(4),\\
p(5)&=p^{\text{col}}(2)\cdot p^{\text{thr}}(1)+p^{\text{col}}(1)\cdot p^{\text{thr}}(3)+p^{\text{col}}(0)\cdot p^{\text{thr}}(5).
\end{aligned} \tag{8}$$

Определяем вероятности отсутствия ошибки аутентификации при разборе дерева сообщений $\overline{p}_i^{\text{col}}(0)$ при числе i сформировавшихся к определённому моменту ветвей дерева по (9):

$$\begin{aligned}
\overline{p}_2^{\text{col}}(0)&=\frac{p^{\text{thr}}(2)p^{\text{col}}(0)}{p^{\text{col}}(1)\cdot p^{\text{thr}}(0)+p^{\text{thr}}(2)p^{\text{col}}(0)},\\
\overline{p}_3^{\text{col}}(0)&=\frac{p^{\text{col}}(0)\cdot p^{\text{thr}}(3)}{p^{\text{col}}(1)\cdot p^{\text{thr}}(1)+p^{\text{col}}(0)\cdot p^{\text{thr}}(3)},\\
\overline{p}_4^{\text{col}}(0)&=\frac{p^{\text{col}}(0)\cdot p^{\text{thr}}(4)}{p^{\text{col}}(2)\cdot p^{\text{thr}}(0)+p^{\text{col}}(1)\cdot p^{\text{thr}}(2)+p^{\text{col}}(0)\cdot p^{\text{thr}}(4)},\\
\overline{p}_4^{\text{col}}(0)&=\frac{p^{\text{col}}(0)\cdot p^{\text{thr}}(4)}{p^{\text{col}}(2)\cdot p^{\text{thr}}(0)+p^{\text{col}}(1)\cdot p^{\text{thr}}(2)+p^{\text{col}}(0)\cdot p^{\text{thr}}(4)},\\
\overline{p}_5^{\text{col}}(0)&=\frac{p^{\text{col}}(0)\cdot p^{\text{thr}}(5)}{p^{\text{col}}(2)\cdot p^{\text{thr}}(1)+p^{\text{col}}(1)\cdot p^{\text{thr}}(3)+p^{\text{col}}(0)\cdot p^{\text{thr}}(5)}.
\end{aligned} \tag{9}$$

Большее число ветвей мы не рассматриваем, так как расчет по приведённым выше формулам (1) в различных диапазонах изменения параметров модели показал, что вероятность формирования 4-х и более отличающихся друг от друга ветвей древовидной структуры в каждой позиции $i = \overline{0\ldots}$ ветви сообщений целевого источника пренебрежимо мала (менее 10^{-3}).

Результаты и их обсуждение. На основании представленных выше формул получены зависимости апостериорной вероятности возникновения ошибки аутентификации от числа сформированных ветвей n^{thr} и длины ветви i, при которой происходит подсчёт числа ветвей (рис. 2). Отличительной особенностью данных графиков является то, что апостериорная вероятность отсутствия ошибок при четырёх побочных ветвях выше, чем при трёх. Причиной этого является то, что четыре ветви возникают тогда, когда есть одна ошибка, порождающая две ветви, и две ветви без ошибок, а вероятность формирования двух ветвей меньше, чем одной, как в случае с тремя ветвями.

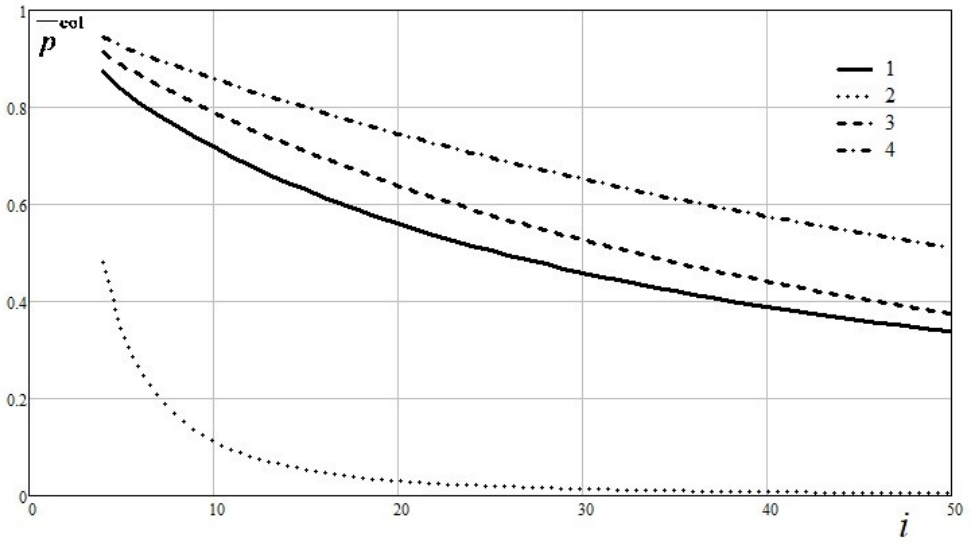


Рисунок 2 – График зависимости апостериорной вероятности $\overline{p}^{\text{col}}$ отсутствия ошибок аутентификации от числа сформированных посторонних ветвей n^{thr} в позиции i цепочки сообщений при U/n = 25: 1) $n^{th\,r}$= 2; 2) n^{thr} = 3; 3) $n^{th\,r}$= 4; 4) n^{thr} = 5

Кроме того, видно, что с ростом значения позиции сообщения априорная вероятность ошибки растёт. При $i > 40$ апостериорная вероятность ошибки при любом количестве посторонних ветвей достигает значений 60–80 %. Это позволяет в таком случае прекращать передачу сообщений от источника и фиксировать возникновение ошибки аутентификации при передаче всей цепочке. Актуальным этот результат является для случаев обмена данными последовательностями, содержащими большое число сообщений. Для исследования ситуаций с небольшой длиной анализируемых ветвей получена зависимость вероятности формирования побочной ветви длиной i от соотношения между длиной поля имитовставки и отношением между числом анализируемых сообщений и длиной цепочки (рис. 3, данные по оси ординат в логарифмическом масштабе).

Данная зависимость хорошо иллюстрирует тот факт, что вероятность формирования побочных ветвей формируемой приёмником древовидной структуры экспоненциально убывает с ростом длины такой ветви. Начиная с i = 5 ею можно пренебречь. Это позволяет сформулировать правило обработки древовидной структуры, заключающееся в следующем: если от какого-либо узла отходит ветвь длиной 6 и более элементов, то все остальные ветви, отходящие от данного узла и имеющие длину меньше 6, можно исключить из обработки. Если ветвей длиной 6 блоков данных обнаруживается более одной, то произошла ошибка формирования цепочки сообщений, при которой невозможно определить цепочку, выданную целевым источником. Дальнейшую передачу сообщений цепочки следует прекратить и перепослать все блоки данных, начиная с того, в котором была обнаружена ошибка.

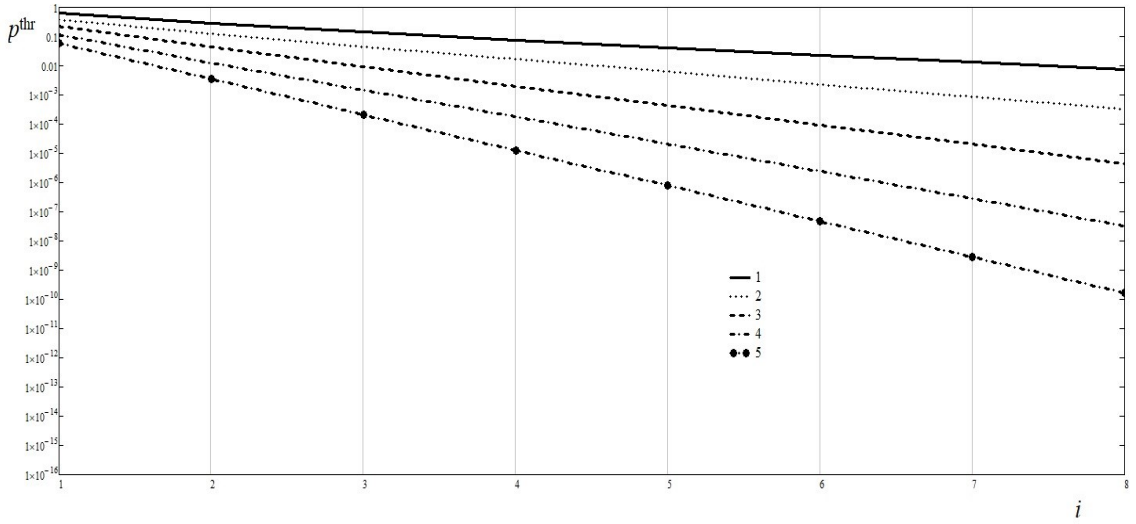


Рисунок 3 – График зависимости вероятности p^{thr} формирования побочной ветви длиной i от соотношения между длиной поля имитовставки и отношением U/n между числом анализируемых сообщений и длиной цепочки: 1) $U/n = 25$; 2) $U/n = 20$; 3) $U/n = 15$; 4) $U/n = 10$; 5) $U/n = 5$

Описанные действия позволят повысить долю полезной информации в общем объёме передаваемых между устройствами данных. Под полезной информацией мы понимаем содержимое всех полей информационного блока, кроме поля имитовставки, которая используется для аутентификации. Данный целевой параметр исчисляется по (10), учитывающей необходимость повторной передачи обработанной с ошибкой цепочки блоков данных [19].

$$K = \frac{L \cdot n \cdot \left(1 - P_{sc}\right)}{\left(L + H\right)\left(n + 2\right)}, \tag{10}$$

где L – длина поля полезной информации блока данных; P_{sc} – априорная вероятность возникновения ошибки аутентентификации источника при передаче цепочки блоков данных.

С учётом того, что P_{sc} при определённых целевых значениях длины поля имитовставки не превышает 0,1, то вероятностью возникновения ошибок в каждом из двух и более последовательно идущих циклов передачи данных можно пренебречь. Считаем, что средний номер блока в цепочке, в котором возникла ошибка $i_{\text{co}\,l} \approx n/2$, и, с учётом уменьшения количества передаваемых блоков в следующем цикле передачи данных, формула (10) для доли полезной информации запишется в виде (11):

$$K' = \frac{L \cdot n}{\left(1 - p^{thr}\right)\sum_{i=0}^{\infty}\left[\left(\frac{n}{2^i} + 2\right)\left(1 - \prod_{j=1}^{\frac{n}{2^i}-1}\left(1 - p_{\text{col}}\left(j\right)\right)^{\frac{n}{2^i}-j}\right)^i \left(L + H + \right]\boldsymbol{log}_2 \frac{n}{2^i}\left[+ 2\right)\right]}, \tag{11}$$

где p^{thr} – вероятность образования побочной ветви из 6 и более блоков.

Результат использования сформулированного выше правила обработки древовидной структуры заключается в увеличении доли полезной информации за счёт уменьшения объёма переспрашиваемой информации (рис. 4).

Из представленного на рисунке 4 графика видно, что в области максимума функции доли полезной информации она дополнительно повышается на 2–5 % в абсолютных значениях и до 30 % в относительных в зависимости от количества обрабатываемых приёмником блоков данных.

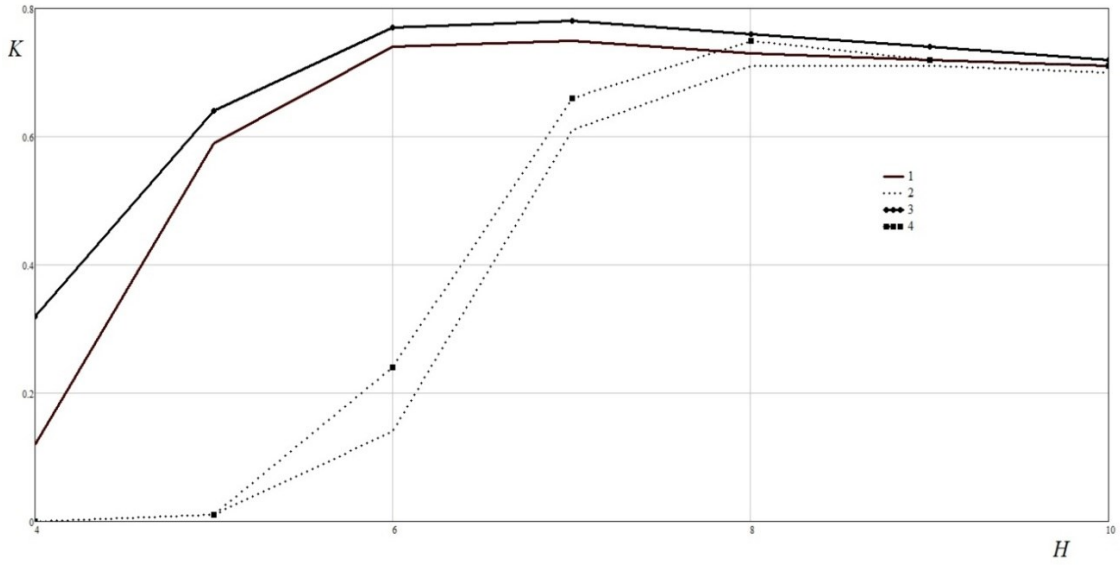


Рисунок 4 – Зависимость отношения K объёму полезной информации к общему объёму обрабатываемой приёмником информации от длины поля H при длине информационной части L = 40 битов, длине цепочки n = 30: 1) $|U|/n$ = 20, без использования правила обработки древовидной структуры; 2) $|U|/n$ = 20, с использованием правила обработки древовидной структуры; 3) $|U|/n$ = 130, без использования правила обработки древовидной структуры; 4) $|U|/n$ = 130, с использованием правила обработки древовидной структуры

Выводы. Результаты исследований показывают, что при использовании методов аутентификации для блоков, зашифрованных в режиме связывания, ошибки аутентификации возникают из-за совпадения содержимого кодов аутентификации сообщений (имитовставок). В результате вместо цепочки сообщений формируется древовидная структура, включающая как основную цепочку блоков, выданных целевым источником, так и некоторое количество посторонних, содержащих информационные блоки, выданные иными источниками. В рамках настоящей работы получены вероятностные оценки формирования посторонних цепочек как в случае правильного выполнения процедуры аутентификации, так и при возникновении ошибки. На основе полученных оценок сформулировано правило обработки ветвей древовидной структуры, подразумевающее исключение из обработки ветвей длиной менее 6. Для сообщений ограниченной длины применение данного правила позволяет увеличить долю полезной информации, обрабатываемой приёмником на 2–5 % за счёт снижения числа переспрашиваемых в результате обнаруженных ошибок аутентификации блоков данных.

Библиографический список

1. Sen, S. Profibus / S. Sen // Fieldbus and Networking in Process Automation. – 2021. – P. 119–144.
2. Sen, S. Modbus and Modbus Plus / S. Sen // Fieldbus and Networking in Process Automation. – 2021. – P. 145–156.
3. Hernández-Vázquez, H. Development of Virtual Router Machine for Modbus Open Connection / H. Hernández-Vázquez, I. Sanchez, F. Martell, J. Guzman, R. Ortiz // Recent Trends in Sustainable Engineering, Proceedings of the 2nd International Conference on Applied Science and Advanced Technology. – Budapest, Hungary, 2021. – P. 1–14.
4. Ondrej, S. ZigBee Technology and Device Design / S. Ondrej, B. Zdenek, F. Petr, H. Ondrej // International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06). – Morne, Mauritius, 2006. – P. 129–129.
5. Knapp, H. RFID systems optimisation through the use of a new RFID network planning algorithm to support the design of receiving gates / H. Knapp, G. Romagnoli // Journal of Intelligent Manufacturing. – 2021.
6. Moon, J. Covert Communications in Time Division Multiple Access Networks / J. Moon // The Journal of Korean Institute of Communications and Information Sciences. – 2021. – Vol. 46 (9). – P. 1407–1410.
7. Dua, A. Covert Communication using Address Resolution Protocol Broadcast Request Messages / A. Dua, V. Jindal, P. Bedi // 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). – Noida, India, 2021. – P. 1–6.
8. Borzenkova, S. Methodology for determining a set of measures to ensure information security in the automated process control system / S. Borzenkova, A. Sychugov // Journal of Physics: Conference Series. – 2020. – Vol. 1679. – P. 1–8.
9. Sviridov, A. The Concept of Information Security in the Process Control System / A. Sviridov, V. Bobkov, D. Bobrikov, A. Balashov // IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). – Saint Petersburg and Moscow, Russia, 2019. – P. 2162–2164.

10. Мальчуков, А. Н. Система автоматизированного проектирования кодеков помехоустойчивых кодов короткой длины / А. Н. Мальчуков, А. Н. Осокин // Известия Томского политехнического университета. – 2008. – Т. 312, № 5. – С. 70–75.

11. Мыцко, Е. А. Исследование алгоритмов вычисления контрольной суммы CRC8 в микропроцессорных системах при дефиците ресурсов / Е. А. Мыцко, А. Н. Мальчуков, С. Д. Иванов // Приборы и системы. Управление, контроль, диагностика. – 2018. – № 6. – С. 22–29.

12. Black, J. CBC MACs for arbitrary-length messages: The three-key constructions / J. Black, P. Rogaway // J. Cryptol. – 2015. – Vol. 18, № 2. – P. 111–131.

13. Liu, C. Implementation of DES Encryption Arithmetic based on FPGA / C. Liu, J. Ji, Z. Liu // AASRI Procedia. – 2013. – Vol. 5. – P. 209–213.

14. Stallings, W. NIST Block Cipher Modes of Operation for Confidentiality / W. Stallings // Cryptologia. – 2010. – № 34 (2). – P. 163–175.

15. Ben Othman, S. An efficient secure data aggregation scheme for wireless sensor networks / S. Ben Othman, H. Alzaid, A. Trad, H. Youssef // IISA 2013. – Piraeus, Greece, 2013. – P. 1–4.

16. Попов, А. Ю. Исследование вариантов реализации алгоритмов Крускала и Прима в вычислительной системе с многими потоками команд и одним потоком данных / А. Ю. Попов // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. – 2015. – № 11. – С. 505–527.

17. Таныгин, М. О. Сложность алгоритма определения источника данных / М. О. Таныгин, Х. Я. Алшаиа, А. В. Митрофанов // Труды МАИ. – 2021. – № 117. – С. 1–21.

18. Таныгин, М. О. Рекурсивный алгоритм формирования структурированных множеств информационных блоков для повышения скорости выполнения процедур определения их источника / М. О. Таныгин, О. Г. Добросердов, Х. Я. А. Алшаиа, В. П. Добрица // Известия Юго-Западного государственного университета. – 2021. – № 2. – С. 51–64.

19. Таныгин, М. О. Теоретические основы идентификации источников информации, передаваемой блоками ограниченного размера : монография / М. О. Таныгин. – Курск : ЗАО «Университетская книга», 2020. – 198 с.

References

1. Sen, S. Profibus. *Fieldbus and Networking in Process Automation*, 2021, pp. 119–144.

2. Sen, S. Modbus and Modbus Plus. *Fieldbus and Networking in Process Automation*, 2021, pp. 145–156.

3. Hernández-Vázquez, H., Sanchez, I., Martell, F., Guzman, J., Ortiz, R. Development of Virtual Router Machine for Modbus Open Connection. *Recent Trends in Sustainable Engineering, Proceedings of the 2nd International Conference on Applied Science and Advanced Technology*. Budapest, Hungary, 2021, pp.1–14.

4. Ondrej, S., Zdenek, B., Petr, F., Ondrej, H. ZigBee Technology and Device Design. *International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06)*. Morne, Mauritius, 2006, pp. 129–129.

5. Knapp, H., Romagnoli, G. RFID systems optimisation through the use of a new RFID network planning algorithm to support the design of receiving gates. *Journal of Intelligent Manufacturing*, 2021.

6. Moon, J. Covert Communications in Time Division Multiple Access Networks. *The Journal of Korean Institute of Communications and Information Sciences*, 2021, vol. 46 (9), 2021, pp. 1407–1410

7. Dua, A., Jindal, V., Bedi, P. Covert Communication using Address Resolution Protocol Broadcast Request Messages. *9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*. Noida, India, 2021, pp. 1–6.

8. Borzenkova, S. Sychugov, A. Methodology for determining a set of measures to ensure information security in the automated process control system. *Journal of Physics: Conference Series*, 2020, vol. 1679, pp. 1–8.

9. Sviridov, A., Bobkov, V., Bobrikov, D., Balashov, A. The Concept of Information Security in the Process Control System. *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. Saint Petersburg and Moscow, Russia, 2019, pp. 2162–2164.

10. Malchukov, A. N., Osokin, A. N. Sistema avtomatizirovannogo proyektirovaniya kodekov pomekhoustoychivykh kodov korotkoy dliny [A computer-aided design system for short-length error-correcting codes]. *Izvestiya Tomskogo politekhnicheskogo universiteta* [Bulletin of the Tomsk Polytechnic University], 2008, vol. 312, no. 5, pp. 70–75.

11. Mytsko, E. A., Malchukov, S. D. Issledovaniye algoritmov vychisleniy kontrolnoy summy CRC8 v mikroprotsessornykh sistemakh pri defitsite resursov [Investigation of algorithms for calculating the CRC8 checksum in microprocessor systems with a shortage of resources]. *Pribory i sistemy. Upravleniye, kontrol, diagnostika* [Devices and systems. Management, control, diagnostics], 2018, no. 6, pp. 22–29.

12. Black, J., Rogaway, P. CBC MACs for arbitrary-length messages: The three-key constructions. *J. Cryptol.*, 2015, vol. 18, no. 2, pp. 111–131.

13. Liu, C., Ji, J., Liu, Z. Implementation of DES Encryption Arithmetic based on FPGA. *AASRI Procedia*, 2013, vol. 5, pp. 209–213.

14. Stallings, W. NIST Block Cipher Modes of Operation for Confidentiality. *Cryptologia*, 2010, no. 34 (2), pp. 163–175.

15. Ben Othman, S., Alzaid, H., Trad, A., Youssef, H. An efficient secure data aggregation scheme for wireless sensor networks. *IISA 2013*. Piraeus, Greece, 2013, pp. 1–4.

16. Popov, A. Yu. Issledovanie variantov realizatsii algoritmov Kruskala i Prima v vychislitelnoy sisteme s mnogimi potokami komand i odnim potokom [Investigation of options for the implementation of the Kruskal and Prima algorithms in a computing system with many command streams and one data stream]. *Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana* [Science and Education: scientific publication of MSTU im. N.E. Bauman], 2015, no. 11, pp. 505–527.

17. Tanygin, M. O., Alshaia, H. Ya., Mitrofanov A. V. Slozhnost algoritma opredeleniya istochnika dannykh [The complexity of the algorithm for determining the data source]. *Trudy MAI* [Proceedings of the MAI], 2021, no. 117, pp. 1–21.

18. Tanygin, M. O., Dobroserdov, O. G., Alshaia, H. Ya. A., Dobritsa, V. P. Rekursivnyy algoritm formirovaniya informatsionnykh blokov dlya povysheniya skorosti vypolneniya protsedur opredeleniya ikh istochnika [Recursive algorithm for the formation of structured sets of information blocks to increase the speed of execution of procedures for determining their source]. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta* [Bulletin of the South-West State University], 2021, no. 2, pp. 51–64.

19. Tanygin, M. O. *Teoreticheskiye osnovy identifikatsii informatsii, peredavaemoy blokami ogranichennogo razmera : monografiya* [Theoretical foundations for identifying sources of information transmitted by blocks of limited size : monograph], Kursk, Publishing House of JSC University Book, 2020. 198 p.

DOI 10.54398/2074-1707_2022_1_93
УДК 681.32; 004.056

СИНТЕЗ СТРУКТУРЫ ОБЪЕКТОВ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ НА ОСНОВЕ МАРКОВСКИХ МОДЕЛЕЙ

Статья поступила в редакцию 29.12.2021, в окончательном варианте – 21.01.2022.

Кочнев Сергей Владимирович, Государственный научно-исследовательский институт приборостроения, 129226, Российская Федерация, г. Москва, пр. Мира, 125,
начальник отдела, ORCID 0000-0001-6226-7518, e-mail: s.v.ko@mail.ru

Лапсарь Алексей Петрович, Управление ФСТЭК России по Южному и Северо-Кавказскому федеральным округам, 344079, Российская Федерация, г. Ростов-на-Дону, ул. Ярослава Галана, 1е/25,
кандидат технических наук, доцент, заместитель начальника отдела, ORCID 0000-0003-2273-725X, e-mail: lapsar1958@mail.ru

Барабошкина Алена Владимировна, Ростовский государственный экономический университет (РИНХ), 344002, Российская Федерация, г. Ростов-на-Дону, ул. Большая Садовая, 69,
магистрант, e-mail: alenaas64@mail.ru.

Отмечается, что безопасная реализация критических производственных процессов в сферах здравоохранения, науки, транспорта, связи, энергетики и других во многом обеспечивается эффективностью функционирования объектов критической информационной инфраструктуры. Указаны проблемы применения информационных технологий в объектах критической информационной инфраструктуры, связанные с повышением их уязвимости к деструктивным воздействиям, реализуемым с помощью удаленного доступа. Установлено, что в настоящее время защита от деструктивных информационных воздействий сводится к прекращению обмена с внешней средой и остановке производственного процесса, что приводит к снижению его эффективности. Проанализированы особенности эксплуатации объектов критической информационной инфраструктуры в условиях деструктивного информационного воздействия и обоснованы требования к ним. Перспективные методы и модели оценки и прогнозирования состояния сложных технических систем распространены на объекты критической информационной инфраструктуры. В качестве информативного параметра в параметризованных моделях объектов информационной инфраструктуры предложено использовать свойства и характеристики деструктивных информационных воздействий. С применением методов структурного синтеза на основе диффузионных марковских моделей разработана обобщенная структура объекта критической информационной инфраструктуры, реализующего диагностические и управляющие функции. Рассмотрен вариант функционирования такого объекта в нормальных и нештатных условиях эксплуатации. Результатами проведенного исследования явилась разработка универсальных требований к объектам критической информационной инфраструктуры, реализующих управление критическими процессами, и обобщенная структура такого объекта. Полученные результаты применимы при разработке перспективных или модернизации существующих объектов критической информационной инфраструктуры критических производственных процессов и реализованы в техническом задании на разработку автоматизированной системы управления субъекта критической информационной инфраструктуры, функционирующего в сфере науки.

Ключевые слова: критический процесс, объект критической информационной инфраструктуры деструктивное информационное воздействие, аварийная ситуация, структура, эволюционные уравнения, параметризованные марковские модели

DESIGNING THE MARKOV MODELS-BASED STRUCTURE FOR CRITICAL INFORMATION INFRASTRUCTURE OBJECTS OF PRODUCTION PROCESSES

The article was received by the editorial board on 29.11.2021, in the final version – 31.01.2022.

Kochnev Sergey V., State Research Institute of Instrument Making, 125 Mira Ave., Moscow, 129226, Russian Federation,
Head of Department, ORCID 0000-0001-6226-7518, e-mail: s.v.ko@mail.ru
Lapsar Alexey P., Management of FSTEC of Russia for the Southern and North Caucasian Federal Districts, 1e/25 Yaroslav Galan St., Rostov-on-Don, 344079, Russian Federation,
Cand. Sci. (Engineering), Associate Professor, Deputy Chief of Department, ORCID 0000-0003-2273-725X, e-mail: lapsar1958@mail.ru
Baraboshkina Alyona V., Rostov State University of Economics (RINH), 69 Bolshaya Sadovaya St., Rostov-on-Don, 344002, Russian Federation,
undergraduate student, e-mail: alenaas64@mail.ru

The secure implementation of critical production processes in the spheres of health, science, transport, communications, energy and other is largely ensured by the effectiveness of critical information infrastructure. The paper indicates problems of applying IT for objects of critical information infrastructure, caused by an increase in their vulnerability to destructive impacts implemented using remote access technologies. The present protection measures are reduced to the termination of exchange with an external environment and a stop of the production process, which leads to a decrease in its effectiveness. The paper analyses operation features of critical information infrastructure objects under destructive information impact conditions and substantiates the requirements for them. Perspective assessment and prediction methods and models for the variety of complex technical systems states are spread to the critical information infrastructure. The properties and characteristics of destructive information influences are proposed the informative parameter in the parameterized models of critical information infrastructure objects. As a result, a generalized structure of the critical information infrastructure object that implements diagnostic and control functions has been developed using structural synthesis methods based on diffusion Markov models. A variant of the functioning of such an object in normal and abnormal operating conditions is considered. The results of the study are the development of universal requirements for information infrastructure objects, which implement management of critical proceedings as well as generalized structure of such objects. The obtained results are applicable in the development of promising or modernization of existing objects of critical production processes and are implemented in the technical specifications for the development of an automated control system of the critical information infrastructure object, which works in the sphere of science.

Keywords: critical process, critical information infrastructure object, destructive information impact, emergency, structure, evolutionary equations, parametrized Markov models

Graphical annotation (Графическая аннотация)

Введение. Нормальное функционирование любого общества предполагает наличие большого числа работоспособных сложных технических систем, обслуживающих самые различные потребности людей, например, в промышленности, энергетике, транспорте, обеспечении жизнедеятельности и т. д. Значительное число таких систем в процессе своей деятельности реализует важнейшие производственные процессы (далее – критические процессы1), нарушение которых может привести к негативным социальным, политическим, экономическим, экологическим последствиям, последствиям для обеспечения обороны страны, безопасности государства и правопорядка. Критические процессы реализуются в соответствии с алгоритмами эксплуатации сложных технических систем в различных условиях обстановки.

[1]Определение критического процесса приведено в «Правилах категорирования объектов критической информационной инфраструктуры Российской Федерации», утвержденных постановлением Правительства Российской Федерации от 8 февраля 2018 года № 12.

Управление критическими процессами осуществляется объектами критической информационной инфраструктуры[1] (КИИ), существенную часть которых можно отнести к значимым, то есть оказывающим существенное влияние на качество жизни и безопасность отдельного человека и общества в целом. К критическим можно отнести процессы функционирования систем жизнеобеспечения крупных населенных пунктов, транспортной инфраструктуры, энергетики, опасные производства и др. Объекты КИИ, обеспечивающие управление критическими процессами в технических системах, являются наиболее чувствительным элементом обеспечения их нормальной реализации. Даже незначительное отклонение от штатной работы управляющего объекта приводит к существенному нарушению критического процесса управляемой технической системы и развитию аварийной ситуации.

На работу и текущее состояние объектов КИИ оказывают существенное влияние условия эксплуатации, внешняя среда, а также различного рода деструктивные воздействия со стороны злоумышленников. При эксплуатации любых технических систем всегда присутствует опасность возникновения аварий, которая с течением времени и выработкой технического ресурса только возрастает [1, 19, 21]. Развитие аварийной ситуации характеризуется дрейфом параметров функционирования, связанных с реализацией критического процесса, за границы области допустимых значений и как следствие – снижением эффективности функционирования технической системы.

Мировой опыт показывает, что в современных условиях постоянно возрастает угроза деструктивных воздействий на объекты КИИ, в том числе и целенаправленных со стороны различного рода злоумышленников, что также приводит к угрозе развития аварийных ситуаций [8, 26, 27]. Специфика функционирования современных объектов КИИ предполагает их взаимодействие с открытыми информационными системами, что привело к существенному возрастанию опасности для них со стороны деструктивных информационных воздействий [4, 7, 25]. Кроме того, созданная на большинстве объектов КИИ система технического обслуживания, связанная с профилактическими мероприятиями по ремонту и замене отдельных блоков, узлов и подсистем, является дорогостоящей и в то же время не гарантирует исправность объекта на период до следующей профилактики.

Очевидно, что использование по предназначению технической системы, реализующей критический процесс, предполагает эффективное управление ее техническим состоянием, которое основывается на достоверной информации о текущем состоянии и призвано учитывать изменение этого состояния в течение некоторого интервала времени [11, 16, 17]. Следовательно, объект КИИ при управления критическим процессом должен решать ряд специфических задач. Во-первых, в обычных условиях повседневной эксплуатации он должен оценивать состояние критического процесса и давать его перспективную оценку, что обеспечит проведение профилактических работ по реальному состоянию как самого объекта КИИ, так и управляемой им технической системы. Во-вторых, в условиях деструктивных воздействий объект КИИ призван исключить возможность развития аварийной ситуации и минимизировать последствия таких воздействий. Очевидно, что и в первом, и во втором случае для реализации оптимального управления критическим процессом необходимо оценить текущее техническое состояние объекта КИИ и сделать перспективную оценку его изменения на некоторый период.

Поскольку при деструктивных воздействиях и связанной с этим угрозы развития аварийной ситуации на первое место выходит скорость принятия управленческих решений и доведения управляющих воздействий, возникает необходимость объединения в рамках объекта КИИ диагностических и управляющих функций [5, 20]. Особую значимость имеют управленческие решения, направленные на купирование деструктивных информационных воздействий, устранение предпосылок к возникновению аварийных ситуаций и минимизацию последствий аварий.

В работах [11, 13] был предложен подход к оценке состояния объектов КИИ с использованием параметризованных уравнений эволюции, позволяющий получать результат в масштабе времени, близком к реальному. Данный подход может служить базой для синтеза перспективных быстродействующих объектов КИИ.

Целью работы является разработка варианта построения структуры объекта КИИ, синтезируемой на основе модифицированного аппарата диффузионных марковских процессов и обеспечивающей эффективное управление критическим процессом как в нормальных условиях эксплуатации, так и в условиях угрозы аварии, спровоцированной деструктивным информационным воздействием.

---

[1] Определение критически важного объекта информационной инфраструктуры приведено в Федеральном законе от 26 июля 2017 года N 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации».

Особенности объектов критической информационной инфраструктуры и требования к ним. Для выбора и обоснования структуры объекта КИИ проанализируем особенности их построения и функционирования, а также требования, предъявляемые к системам диагностики и управления.

Объекты КИИ включают в себя совокупность различных элементов, блоков, узлов и подсистем, отличающихся по принципам функционирования, целевому назначению, техническими и эксплуатационными характеристиками. То есть объекты КИИ обладают признаками сложных систем, что требует специфических подходов к их синтезу. Выбор модели функционирования объекта КИИ как сложной системы, которая служит базой для синтеза его структуры, предполагает учет следующих особенностей:

1. Сложный и неоднозначный алгоритм взаимодействия составных частей и узлов объекта КИИ в процессе реализации заданных функций.

2. Непрерывный режим работы объекта КИИ при необходимости обеспечения его эффективности не ниже заданной, особенно в условиях нештатной ситуации.

3. Уникальность большинства объектов КИИ, что затрудняет набор статистических данных для оценки их параметров функционирования в процессе эксплуатации.

4. Объекты КИИ при отказах не обязательно теряют работоспособность, отказы могут приводить только к изменению их характеристик.

5. Обеспечение высокой надежности объектов КИИ обеспечивается наличием функциональной (информационной, алгоритмической, другой) избыточности, а также резервированием ключевых элементов.

6. На объектах КИИ используются сложные разнородные системы диагностики, что, в свою очередь, порождает проблему надежности этих систем и компенсацию ошибок первого и второго рода.

7. Субъективный фактор при эксплуатации объекта КИИ, роль персонала, обеспечивающего критический процесс, резко возрастает в условиях дефицита времени, связанного с возникновением и развитием аварийной ситуации.

Для обеспечения эффективности оценки состояния технической системы и формирования управления критическим процессом требуется выбрать или синтезировать адекватную модель объекта КИИ. Известно [6, 20], что система является сложной, если она включает в себя большое количество взаимодействующих элементов (узлов, блоков, подсистем) и способна реализовывать сложный алгоритм функционирования. Данное определение позволяет отнести объекты КИИ к категории сложных систем. Объект КИИ как сложная система при выходе из строя отдельных элементов составных частей не теряет работоспособность, зачастую только снижаются характеристики ее эффективности. Отказ объекта КИИ заключается в дрейфе параметров его функционирования и характеристик эффективности за пределы установленных допусков [1, 10, 20, 27].

Для формального описания текущего состояния объектов, аналогичных объектам КИИ, широко используются модели на основе диффузионных марковских процессов [3, 11, 18]. Однако для эффективного использования при синтезе объектов КИИ, подверженных деструктивным информационным воздействиям, они требуют некоторой доработки. То есть необходимо применение предложенного в работах [11, 13] модифицированного марковско-параметрического аппарата оценки состояния сложных технических систем для синтеза на его основе объектов КИИ. Анализ инцидентов на объектах КИИ, произошедших в мире за последние несколько лет, показывает, что деструктивное воздействие, осуществляемое путем дистанционного вмешательства в системы управления объектов через информационную сферу, происходит значительно чаще, чем традиционное физическое воздействие вследствие антропогенных, технических или природных факторов [4, 7, 15, 23]. При дальнейшем рассмотрении функционирования объектов КИИ будем считать, что наиболее вероятным воздействием на них является деструктивное информационное воздействие (ИВ), осуществляемое путем удаленного доступа через информационно-телекоммуникационные каналы. Таким образом, синтезируемый объект КИИ должен эффективно функционировать в условиях угрозы и реализации данного воздействия.

Широко применяемые в настоящее время на технических объектах системы диагностики и управления используют в основном методы и модели теории надежности [5, 10, 26]. Такие системы недостаточно эффективны в условиях, связанных с возникновением аварийной ситуации. Аварийная ситуация, спровоцированная деструктивным информационным воздействием, накладывает специфические особенности на функционирование объекта КИИ:

1. Существенный дефицит времени на оценку текущего технического состояния и выработку управляющих воздействий по парированию деструктивного воздействия и минимизации последствий аварии.

2. Снижение достоверности оценки текущего состояния объекта КИИ, связанное с ростом стохастичности исследуемых параметров, вызванного повышением уровня и интенсивности внутренних шумов и внешних помех.

3. Некоторые параметры, характеризующие техническое состояние объекта КИИ, могут выходить за границы области допустимых значений без потери работоспособности объекта. Это обусловливает рост параметрической неопределенности, необходимость расширения диапазонов измерений и, как следствие, приводит к лавинообразному увеличению вычислительных затрат при решении задач диагностики.

4. Преобладание динамической погрешности измерений из-за специфики поведения измеряемых величин в аварийной ситуации.

5. Возможность развития аварийной ситуации или усугубления ее последствий из-за неоптимальных управленческих решений.

С учетом указанной специфики реализации критического процесса в условиях аварийной ситуации задачи диагностики и управления предполагают создание более совершенных объектов КИИ с элементами искусственного интеллекта и функциями экспертных систем, основанных на эффективных моделях оценки технического состояния и выработки квазиоптимальных управляющих воздействий.

На основании сказанного выше сформулируем основные требования к синтезируемому объекту КИИ. Объект КИИ должен обеспечить:

- непрерывный мониторинг технического состояния с целью выявления деструктивного воздействия и предпосылок появления аварийной ситуации;
- возможность функционирования как в штатных условиях эксплуатации, так и в условиях аварийной ситуации, вызванной деструктивным воздействием;
- высокое быстродействие средств измерений и системы диагностики технического состояния;
- высокую оперативность выработки рекомендаций на формирование управляющих воздействий и доведения их до исполнительных устройств системы управления;
- требуемую точность и достоверность оценки технического состояния объекта КИИ;
- коррекцию области допустимых значений параметров в зависимости от решаемой задачи и уровня шумов и помех;
- использование для задач диагностики и управления набора вещественных параметров, влияющих на функционирование объекта КИИ, которые оцениваются специально созданной системой оценки характеристик воздействия или назначаются системой высшего уровня;
- возможность использования априорных сведений о техническом состоянии и поведении параметров объекта КИИ;
- принципиальную возможность использования на любых объектах КИИ, независимо от назначения и ведомственной принадлежности, после незначительной доработки и привязки к конкретному объекту;
- возможность адаптации объектов КИИ к условиям реализации критического процесса: корректировку модели объекта, области допустимых значений и диапазонов измерения, требуемой точности, выбор наиболее информативных параметров и т. д.;
- выработку управляющих воздействий и оценку последствий реализации управления критическим процессом;
- разработку рекомендаций для обслуживающего персонала по дальнейшему использованию объекта КИИ после купирования деструктивного воздействия и ликвидации аварии.

Особенности построения и функционирования объекта КИИ, недостатки существующих автоматизированных систем контроля и управления технологическими процессами, предъявляемые к объекту КИИ требования, результаты разработки параметризованных марковских моделей [6, 11, 13] позволяют синтезировать возможный вариант структуры объекта КИИ, управляющего критическим процессом с использованием модифицированного аппарата диффузионных марковских процессов.

Структура объекта КИИ. Возможный вариант структуры объекта КИИ, основанной на параметризованных марковских моделях и обеспечивающей реализацию приведенных выше требований, представлен на рисунке 1.

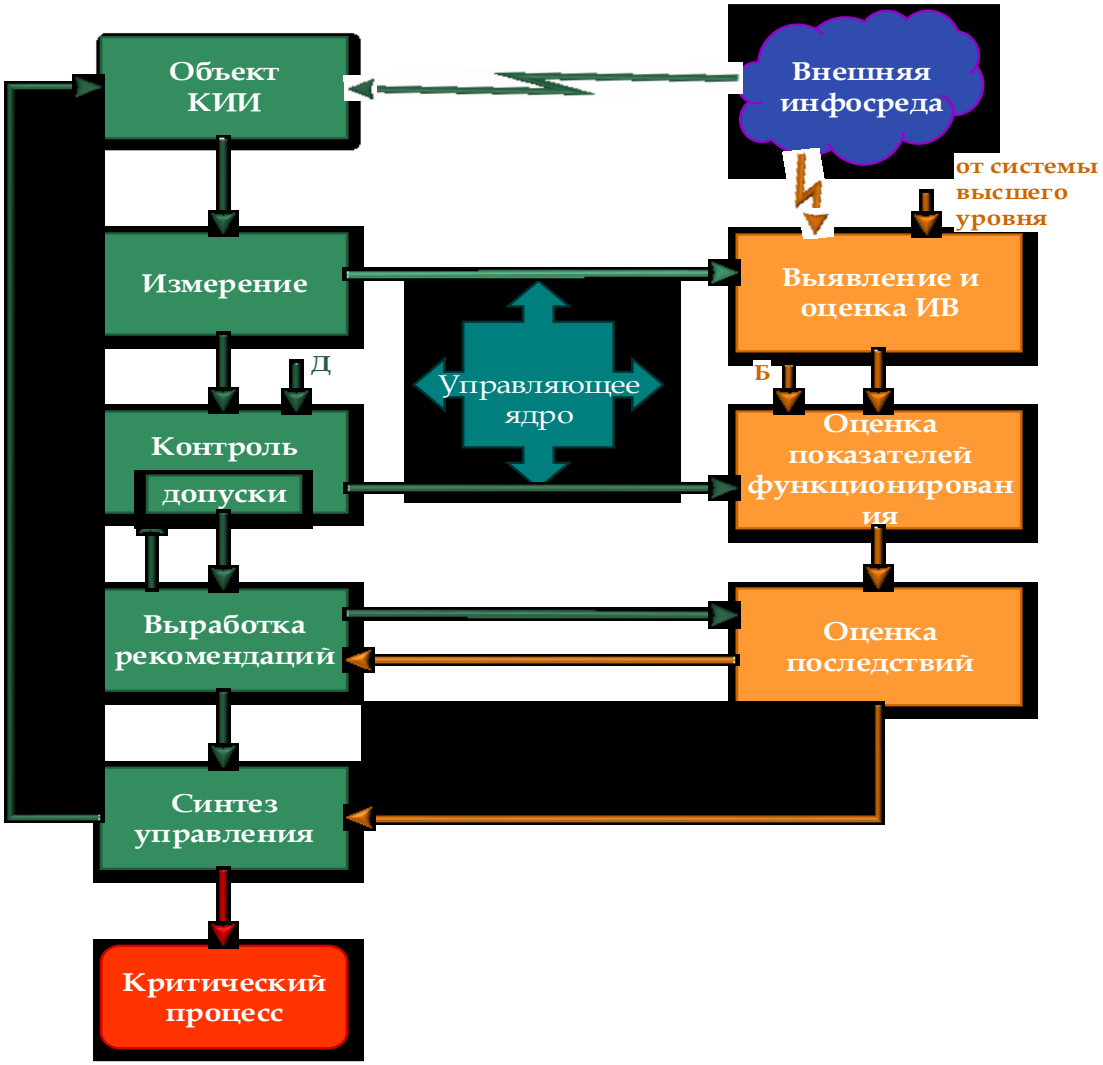


Рисунок 1 – Структура объекта КИИ

Конструктивно система включает две составные части: диагностическую и управляющую. Диагностическая часть объекта КИИ имеет ряд особенностей, позволяющих обеспечить выполнение указанных выше требований. Она включает в себя две составляющих: собственно измерительную и логическую, которая реализует алгоритмы диагностики и оценки технического состояния объекта КИИ, а также выполняет функции экспертных систем поддержки решений персонала при управлении критическим процессом. Нужно отметить, что провести четкую границу между названными составными частями не всегда представляется возможным, так как отдельные блоки выполняют функции, присущие обеим компонентам.

Управляющая подсистема в части построения исполнительных устройств разрабатывается для каждого критического процесса и в данной работе не рассматривается.

Синтезируемый объект КИИ функционально содержит два относительно независимых канала, один из которых функционирует в любых условиях эксплуатации (на рисунке показан зеленым цветом), а второй (выделен бежевым цветом) только в условиях деструктивного информационного воздействия.

Основным элементом всего объекта КИИ является управляющее ядро, представляющее собой микропроцессор и запоминающее устройство, в котором содержится база знаний и правил, обеспечивающая реализацию алгоритмов функционирования объекта КИИ. Управляющее ядро в соответствии с правилами, изложенными в [11–13], формирует марковскую математическую модель объекта КИИ вида $\frac{\partial p(x,\omega,t)}{\partial t} = -a(x,\omega,t)\frac{\partial p(x,\omega,t)}{\partial x} + \frac{1}{2}b(x,\omega,t)\frac{\partial^2 p(x,\omega,t)}{\partial x^2}$, а также содержит базовые решения параметризованных эволюционных уравнений.

Запоминающее устройство содержит алгоритмы обработки информации, варианты возможных решений и рекомендаций обслуживающему персоналу, признаки состояния объекта, другую информацию, необходимую для обеспечения эффективного функционирования. Кроме того, в базе знаний хранится и постоянно обновляется статистическая информация о состоянии объекта в течение всего времени его эксплуатации в виде параметров функционирования и обобщенных стохастических характеристик объекта. Управляющее ядро позволяет функционировать объекту КИИ как интеллектуальной экспертной системе [2, 9, 14, 22].

Диагностическая подсистема включает в себя блок измерений, содержащий разнообразные датчики, измерительные преобразователи, приборы, другие средства измерений и блок контроля. Блок измерений синтезируемого объекта КИИ призван не только получать измерительную информацию о характеристиках управляемого критического процесса, но и по заданным критериям выявлять наличие деструктивного информационного воздействия. На основе анализа измерительной информации в блоке контроля с учетом назначенных допусков и ограничений на характеристики осуществляется оценка и анализ текущего состояния объекта КИИ.

Первичная измерительная информация, результаты текущего контроля, марковская модель объекта КИИ, синтезированная в виде параметризованных дифференциальных (эволюционных) уравнений и размещенная в базе знаний и правил, а также соответствующие алгоритмы позволяют получить значения основных показателей функционирования исследуемого объекта, что реализуется в соответствующем блоке [3, 11–13].

Предварительно синтезированные опорные решения хранятся в базе знаний и правил, постоянно уточняются по мере накопления статистической информации, до момента появления признака аварийной ситуации, вызванной деструктивным ИВ, факт наличия и характеристики которого поступают от соответствующего блока [11, 16, 25].

Признаком аварийной ситуации может служить пересечение информативным параметром границы области допустимых значений или существенное возрастание скорости его изменения. Кроме того, признак аварийной ситуации или присутствие деструктивного ИВ могут поступить от системы высшего уровня. После оценки основных показателей функционирования объекта в блоке выработки рекомендаций осуществляется формирование допустимых вариантов квазиоптимальных управленческих воздействий на критический процесс с целью купирования деструктивного ИВ и недопущения развития аварийной ситуации. Далее производится оценка возможных последствий реализации каждого варианта управленческого воздействия. После выбора наиболее приемлемого варианта соответствующий блок осуществляет синтез управляющих воздействий и доведение их до объекта. Доведение воздействия может производиться по команде обслуживающего персонала, либо, в условиях дефицита времени, в автоматическом режиме.

Как отмечалось выше, функционирование объекта КИИ может происходить как в нормальных условиях, так и в условиях аварийной ситуации. В зависимости от условий будут отличаться показатели эффективности функционирования объекта КИИ. В нормальных условиях основное требование к объекту КИИ – это высокая точность и достоверность оценки текущего состояния объекта, а в условиях аварийной ситуации – как можно меньшее время получения результата оценки состояния при заданной точности, т. е. быстродействие. Таким образом, оптимизация функционирования объекта КИИ должна осуществляться на основании критерия точности и достоверности «Д» в нормальных условиях и критерия быстродействия «Б» в условиях аварийной ситуации. Названные критерии могут назначаться обслуживающим персоналом объекта, системой высшего уровня, либо формироваться самим объектом КИИ исходя из складывающейся ситуации.

Для повышения эффективности алгоритм синтезируемого объекта КИИ функционирования должен обеспечивать адаптацию к условиям решения задач диагностики и управления [20, 24]. При нормальных условиях функционирования объекта КИИ точность измерений может быть значительно выше, чем при измерениях в условиях деструктивного ИВ и развития аварийной ситуации, которая предполагает существенное снижение времени на принятие управленческого решения. Адаптацией считается также изменение уровня допусков (области допустимых значений параметров) в условиях выявления деструктивного ИВ. Функция адаптации реализовывается преимущественно программно-алгоритмическими методами, отличающимися высоким быстродействием и низкой стоимостью по сравнению с техническими.

Рассмотрим в общем виде назначение основных элементов структуры объекта КИИ.

Техническое состояние объекта КИИ характеризуется соответствующим вектором $X(\omega,t)$, где $\omega \in \Omega$ – вещественный параметр, оказывающий существенное влияние на процессы эксплуатации объекта, в рассматриваемом случае – это характеристики деструктивного информационного воздействия: сложность, интенсивность, продолжительность и др.

Блок измерений выполняет ряд функций:

- осуществляет с заданной точностью измерение характеристик объекта КИИ и управляемого им критического процесса $x_i(t)$, $\overline{i=1,N}$;
- осуществляет первичную обработку измерительной информации и оценку вектора технического состояния $X(\omega,t)=F_{\text{È}}(X(\omega,t))$;
- оценивает параметр $\omega \in \Omega$ и определяет его основные свойства;
- определяет признак (признаки) возникновения аварийной ситуации либо по значению информативного параметра, $\Psi=[F_{\text{Èi}}(x_i(\omega,\text{t})) \notin x_{\text{äîï}}]$, $i=\overline{1,n}$, либо по скорости его изменения $\Psi = F_{\text{Èi}}\left[\left(\frac{\partial x_i}{\partial t}\right) \geq \left(\frac{\partial x_i}{\partial t}\right)_{\text{äîï}}\right]$.

Управляющее ядро координирует работу всех блоков, управляет процессами контроля и выработки решений, взаимодействует с обслуживающим персоналом и, при необходимости, с системой высшего уровня. Процессор на основе моделей, хранящихся в базе знаний и правил управляющего ядра, вычисляет локальные характеристики марковского процесса, моделирующего техническое состояние объекта КИИ (коэффициенты сноса и диффузии), и синтезирует модель объекта $Y=F_{\text{ÁÄ}}(X_{\text{ñò}}(\omega,t_{\text{íàáë}}),X(\omega,t))$ с учетом адаптации системы к текущим условиям эксплуатации на основе критериев достоверности и быстродействия [11, 12, 24]. При вычислении локальных характеристик используется не только текущее значение вектора состояния объекта, но и статистические данные за все время наблюдения $X_{\text{ñò}}(\omega,t_{\text{íàáë}})$, которые хранятся в запоминающем устройстве управляющего ядра. Математическая модель объекта КИИ строится путем синтеза параметризованных эволюционных уравнений вида $D=F_{\text{ÄÓ}}(X_{\text{ñò}}(\omega,\text{Ò}_{\text{íàáë}}),X(\omega,t),Y(\omega,t))$, где $\text{Ò}_{\text{íàáë}}$ – время наблюдения за объектом с момента начала его эксплуатации [11, 18], на основе текущего значения вектора состояния $X(\omega,t)$, статистической информации о состоянии объекта за все время эксплуатации, хранящейся в базе, и результатов контроля состояния объекта $Y(\omega,t)=F_K(X(\omega,t))$. Далее вычисляются базовые решения эволюционных уравнений, моделирующих объект КИИ, сразу в параметрическом, т. е. зависящем от вектора параметров ω виде. По мере поступления дополнительной информации эти решения уточняются $Q=F_{\text{ÄÓ}}(L(\omega,t),K_{\text{Ò}})$. Базовые решения эволюционных уравнений используются при вычислении показателей функционирования объекта как в нормальных условиях, так и в случае возникновения признаков возникновения аварийной ситуации, вызванной деструктивным ИВ, $Q=F_{\text{ÄÓ}}(Q(\omega,\text{t}),t_{\text{íàáë}},\Psi)$ [11, 13].

Оценка показателей функционирования с использованием опорных решений параметризованных эволюционных уравнений и прогнозирование их поведения в последующие моменты времени осуществляется в виде функционала $Z=F_{\text{ÏÔ}}(P(\omega,t),\omega)|_{\text{Á,Ä}}$. Качество оценки показателей функционирования определяется критерием достоверности «Д» при нормальной эксплуатации и критерием быстродействия «Б» при возникновении признака аварийной ситуации. Номенклатура критериев достоверности «Д» и быстродействия «Б» формируется при проектировании объекта КИИ применительно к каждому критическому процессу. Назначение же необходимого критерия осуществляется по хранящемуся в базе знаний и правил заданному алгоритму. Основанием для выбора конкретного критерия служит анализ текущего состояния объекта и наличие признака деструктивного воздействия $\Psi(\omega)$.

Адаптация реализуется в виде функции $F_i^A(\omega,t)=A_j(F_i(\omega,t))$, где A_j – j-й вид адаптации. При измерении и контроле адаптация заключается в изменении точности, чувствительности и быстродействия, смене диапазонов входящих в него средств измерений в зависимости от условий функционирования объекта КИИ. При оценке показателей функционирования объекта КИИ адаптация реализуется в виде выбора метода оценки локальных характеристик моделируемого параметризованного эволюционного уравнения на основе совокупности имеющейся информации, формирования набора локальных характеристик, наиболее важных для решаемой задачи, и точности их оценки; назначения требуемого набора показателей функционирования объекта исходя из условий и установления точность их оценки. При синтезе управления адаптируется вид и уровень

управляющего воздействия на объекте КИИ с целью ликвидации аварийной ситуации, обеспечения нормальной эксплуатации и оптимального функционирования объекта.

Блок выработки рекомендаций сравнивает полученные показатели функционирования и прогноз их изменения с набором допустимых значений, а затем формирует виды и уровни возможных воздействий на объект [1, 2, 16] для восстановления его нормального функционирования $B(x,\omega) = F_{\text{ÁÐ}}(C(x,\omega))|_{\text{Ä},\text{Á}}$.

Блок оценки возможных последствий выбранного воздействия на объект моделирует поведение объекта КИИ после реализации возможных воздействий на него по устранению аварийной ситуации или последствий аварии $X(\omega,t_{\text{ÀÑ}}) = F_{\text{ÀÏ}}\left(X(\omega,t),U(\omega,t),t\right)$, оценивает (прогнозирует) последствия воздействия и выбирает оптимальные вид и уровень $B^*(x,\omega)$ воздействий из всей совокупности возможных видов воздействия $B(x,\omega)$. Работа данного блока происходит с учетом общих критериев «Д» и «Б» и функционала качества

$$J = M\left\{X^T(T \mid X_0,\omega)GX(T \mid X_0,\omega) + \right.$$

$$\left. + \int_0^T \left[X^T(t \mid X_0,\omega)Q(t)X(t \mid X_0,\omega) + U^T(t,X \mid X_0,\omega)U(t,X \mid X_0,\omega)\right]dt\right\},$$

где $U = U(\omega,t,X \mid X_0) \in R^k$ – вектор управляющих воздействий, минимизирующий приведенный выше функционал, $G \in R^{r\times r}$, $Q(t) \in R^{r\times r}$ – заданные положительно определенные матрицы.

И, наконец, при синтезе управления реализуется выбранное квазиоптимальное управляющее воздействие $U = F_{\text{Ó}}\left(X(\omega,t),Z(x,\omega),R(x,\omega)\right)$ и доводится до исполнительных механизмов и устройств на объекте управления, которым является критический процесс.

Функционирование объекта КИИ. Функционирование объекта КИИ будем рассматривать во время реализации управления критическим процессом. Считаем, что в текущий момент времени проведены подготовительные мероприятия: определены коэффициенты сноса и диффузии эволюционных уравнений (локальные характеристики модели исследуемого объекта КИИ) и сформированы аналитико-параметрические решения соответствующих диффузионных уравнений. В базе данных хранятся опорные решения диффузионных уравнений, правила и алгоритмы для оценки состояния объекта КИИ, а также варианты возможных рекомендаций для синтеза управления критическим процессом в различных условиях и другая информация, необходимая для обеспечения устойчивого функционирования объекта КИИ и управляемого им критического процесса. Управляющее ядро на базе анализа дополнительной (текущей) измерительной информации уточняет опорные решения и корректирует алгоритмы выработки рекомендаций. Объект КИИ реализует свои функции как при нормальных условиях эксплуатации, так и при возникновении предпосылок к развитию аварийной ситуации, вызванной деструктивным ИВ.

В нормальных условиях эксплуатации объект КИИ обеспечивает оптимальность реализации критического процесса в соответствии с его предназначением. Определяющим показателем качества функционирования объекта КИИ в этих условиях является точность, характеризуемая критерием «Д». При этом должны быть соблюдены ограничения на скорость получения измерительной информации, определяемую критерием «Б». Ограничения обусловлены быстродействием средств измерений, входящих в состав объекта КИИ, и необходимостью обеспечения измерения параметров функционирования объекта в статическом режиме. В условиях нормального функционирования объект КИИ накапливает статистическую информацию для корректировки математической модели. При нормальных условиях эксплуатации объект КИИ, в зависимости от конкретных значений вектора параметров $\omega \in \Omega$ с использованием опорных решений, оценивает необходимые показатели функционирования [11, 13, 17]. Такое рассмотрение задачи оценки показателей функционирования объекта позволяет перенести основные временные затраты на этап подготовительных мероприятий и тем самым существенно повысить быстродействие объекта КИИ [11]. Полученная оценка показателей функционирования служит базой для выработки рекомендаций по поддержанию объекта в работоспособном состоянии с последующим синтезом управления.

Метод синтеза параметризованных эволюционных уравнений, моделирующих объект КИИ, на основе априорной информации описан в работах [11–13]. Решения названных уравнений синтезируются сразу в параметрическом (зависящем от вектора параметров ω) виде. Полученные параметризованные решения являются опорными, хранятся в базе данных управляющего ядра и используются для оперативной выработки управляющих воздействий при обнаружении деструктивно-

го воздействия и возникновении аварийной ситуации. Уточнение модели объекта и опорных решений производится в процессе штатной эксплуатации по мере накопления статистической информации. Адаптация с учетом выбранного критерия осуществляется путем изменения передаточных функций измерительных устройств, блоков оценки показателей функционирования критического процесса и формирования управляющего воздействия. Объект КИИ в нормальных условиях реализует функцию прогноза основных параметров функционирования и технического состояния (показателей функционирования) объекта. Прогноз названных показателей обеспечивает оптимальный режим функционирования объекта КИИ в соответствии с его предназначением, а также позволяет оптимизировать процессы эксплуатации, в частности – определить периодичность и объем технического обслуживания и ремонта. Это позволяет сократить затраты на техническое обслуживание и устранение неисправностей, снизить вероятность возникновения внезапных отказов.

Особенности функционирования объекта КИИ, вызванные деструктивным ИВ и последующим развитием аварийной ситуации, связаны с дефицитом времени на определение текущего технического состояния, оценку его изменения во времени и принятием управленческого решения. Признак деструктивного воздействия может поступать от обслуживающего персонала, системы высшего уровня или выявляется соответствующими измерительными системами и транслируется в блок, осуществляющий оценку основных свойств и характеристик деструктивного информационного воздействия. Для выявления деструктивного воздействия в объекте КИИ может создаваться отдельный канал. Одновременно в качестве приоритетного вводится критерий быстродействия «Б», изменяя тем самым основной показатель качества функционирования объекта КИИ. В этом режиме качество объекта КИИ определяется быстротой выработки и доведения до исполнительных устройств управляющих воздействий при заданных ограничениях на точность. Ограничения связаны с достоверностью оценки вектора технического состояния, т. е. значение критерия точности «Д» должно обеспечить минимизацию ошибок первого и второго рода.

Определив значение параметра $\omega \in \Omega$ (измерив или получив от смежной системы) и хранящиеся в базе базовые решения эволюционных уравнений, оперативно вычисляются показатели функционирования объекта, на основе которых объект КИИ выдает рекомендации по формированию управления критическим процессом. Далее происходит оценивание (моделирование) возможных последствий их реализации, и в случае неудовлетворительного результата последствий выработанных рекомендаций принимаются новые рекомендации [11, 12, 17]. Оценка последствий реализации рекомендаций производится на базе заранее сформированного набора возможных аварийных ситуаций и модели как самого объекта КИИ, так и управляемого критического процесса. Синтезированное на основе выбранного решения квазиоптимальное управляющее воздействие доводится до соответствующего критического процесса.

Заключение. В работе проанализированы основные свойства объекта КИИ и на этой основе сформулированы основные требования к системам диагностики и управления таких объектов. Рассмотрены особенности функционирования объекта КИИ в условиях развития аварийной ситуации, обусловленной деструктивным информационным воздействием на объект. Для моделирования объекта КИИ предложено использование аппарата параметризованных диффузионных марковских процессов. Синтезирован вариант структурной схемы объекта КИИ, обеспечивающий реализацию предъявленных требований, рассмотрена ее работа в различных условиях функционирования объекта.

Предложенный в работе подход к синтезу структуры применим для перспективных объектов КИИ, осуществляющих управление критическими процессами. Синтезированная в ходе исследования структура объекта КИИ предназначена для работы как в нормальных условиях, так и при возникновении аварийной ситуации, связанной с деструктивными воздействиями на объект и выходом параметров функционирования за пределы области допустимых значений. Реализация разработанного объекта КИИ в штатных условиях эксплуатации дает возможность оптимизации основных эксплуатационных процессов, в том числе объема и времени профилактических мероприятий на объекте.

Максимальный эффект от внедрения разрабатываемой системы следует ожидать при угрозе возникновения аварийной ситуации. При деструктивном воздействии объект КИИ позволяет получать оперативную оценку показателей функционирования на основе определения значения информативного параметра в реальном масштабе времени. В этих условиях синтезированная система обеспечивает повышение безопасности критического процесса за счет снижения вероятности принятия персоналом ошибочных управленческих решений при минимуме времени на оценку обстановки.

Предлагаемый подход может быть использован для формирования технических требований к автоматизированным системам управления техническими процессами вновь создаваемых и модернизируемых технических систем субъектов КИИ. Указанные в работе особенности эксплуата-

ции объектов КИИ и основные требования к ним учитывались в Государственном научно-исследовательском институте приборостроения при составлении технического задания на модернизацию такой системы.

Объекты КИИ с предложенной в работе структурой применимы в любых сложных экономических, социальных, экологических и им подобным системах, поведение которых возможно смоделировать уравнениями эволюции.

Библиографический список

1. Абзалов, А. В. Методика анализа предаварийных ситуаций на технологических объектах управления / А. В. Абзалов, Р. Р. Жедунов // Прикаспийский журнал: управление и высокие технологии. – 2013. – № 4. – С. 50–59.

2. Аль-Бусаиди, С. С. С. К вопросу о поддержке процесса принятия решения об улучшении деятельности в испытательной лаборатории / С. С. С. Аль-Бусаиди, Ю. Н. Воякина, С. В. Пономарев // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 1. – С. 27–45.

3. Артеменков, С. Л. Марковские модели в задачах диагностики и прогнозирования / С. Л. Артеменков, В. И. Алхимов, С. Н. Баранов, О. Б. Беляева, П. Н. Думин, П. А. Корниенко, Л. С. Куравский, С. Б. Малых, А. А. Марголис, П. А. Мармалюк, А. С. Панфилова, С. И. Попков, Г. А. Юрьев, Н. Е. Юрьева ; под ред. Л. С. Куравского. – Москва : Изд-во МГППУ, 2017. – 203 с.

4. Бачманов, Д. А. Исследование вопросов совершенствования систем защиты от DDos-атак на основе комплексного анализа современных механизмов противодействия / Д. А. Бачманов, А. Р. Очередько, М. М. Путято, А. С. Макарян // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 1. – С. 63–74.

5. Бугайченко, П. Ю. Динамическая модель оценивания качества подготовки и применения сложной технической системы / П. Ю. Бугайченко, А. И. Данилов, А. М. Зубачев // Труды военно-космической академии им. А.Ф. Можайского. – 2018. – № 664. – С. 20–26.

6. Бункин, Н. Ф. Стохастические системы в физике и технике / Н. Ф. Бункин. – Москва : МГТУ им. Баумана. – 2011. – 366 с.

7. Володин, И. В. Классификация механизмов атак и исследование методов защиты систем с использованием алгоритмов машинного обучения и искусственного интеллекта / И. В. Володин, М. М. Путято, А. С. Макарян, В. Ю. Евглевский // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 2. – С. 91–98.

8. Данилов, А. И. Методика численного анализа эффективности подготовки и применения сложной технической системы / А. И. Данилов, А. М. Зубачев, А. А. Данилов // Известия тульского государственного университета. Технические науки. – 2018. – № 8. – С. 186–199.

9. Ирадж, Эльяси Комари. Анализ задач разработки и реинжиниринга компьютерных сетей для критических приложений / Ирадж Эльяси Комари, А. В. Горбенко //Радюелектронт і компютерш системи. – 2006. – № 7 (19). – С. 32–35.

10. Кляцкин, В. И. Очерки по динамике стохастических систем / В. И. Кляцкин. – Москва : Крассанд, 2012. – 448 с.

11. Кочнев, С. В. Параметрический подход к оценке состояния сложных технических объектов на основе эволюционных уравнений / С. В. Кочнев, А. П. Лапсарь // Прикаспийский журнал: управление и высокие технологии. – 2017. – № 2. – С. 10–18.

12. Кубарев, А. В. Параметрическое моделирование состояния объектов критической инфраструктуры в условиях деструктивного воздействия / А. В. Кубарев, А. П. Лапсарь, С. А. Назарян // Вопросы кибербезопасности. – 2021. – № 3. – С. 58–67.

13. Кубарев, А. В. Повышение безопасности эксплуатации значимых объектов критической инфраструктуры с использованием параметрических моделей эволюции / А. В. Кубарев, А. П. Лапсарь, Я. В. Федорова // Вопросы кибербезопасности. – 2020 – № 1. – С. 8–17.

14. Курейчик, В. М. Особенности построения систем поддержки принятия решений / В. М. Курейчик // Известия Южного федерального университета. Технические науки. – 2012. – Тематический выпуск. – С. 92–98.

15. Максименко, В. Н. Основные подходы к анализу и оценке рисков информационной безопасности / В. Н. Максименко, Е. В. Ясюк // Экономика и качество систем связи. – 2017. – № 2. – С. 42–48.

16. Михеев, М. Ю. Математические и информационно-структурные модели прогнозирования состояния технически сложных объектов / М. Ю. Михеев, О. В. Прокофьев, А. Е. Савочкин, М. А. Линкова // Прикаспийский журнал: управление и высокие технологии. – 2015. – № 4. – С. 232–249.

17. Николаев, А. Б. Автоматизация поддержки управленческих решений при организации наукоемкого производства на основе гибкой обратной связи / А. Б. Николаев, П. С. Якунин, Ю. А. Краснов // Московский автомобильно-дорожный государственный технический университет. – 2013. – № 4. – С. 3–13.

18. Оксендаль, Б. Стохастические дифференциальные уравнения. Введение в теорию и приложения / Б. Оксендаль. – Москва : Мир ; АСТ, 2003. – 408 с.

19. Острейковский, В. А. Вероятностное прогнозирование работоспособности элементов ЯЭУ / В. А. Острейковский, Н. Л. Сальников. – Москва : Энергоатомиздат, 1990. – 416 с.

20. Пугачев, В. С. Стохастические дифференциальные системы / В. С. Пугачев, И. Н. Синицын. – Москва : Наука, 2000. – 1000 с.

21. Савочкин, А. Е. Алгоритмизация работы системы мониторинга и контроля для решения задач идентификации степени повреждения технически сложных объектов / А. Е. Савочкин // Прикаспийский журнал: управление и высокие технологии. – 2014. – № 2. – С. 23–35.

22. Сибиряков, М. А. Модификация и моделирование алгоритмов обработки данных в кэш-памяти систем хранения данных / М. А. Сибиряков, Е. С. Васяева // Кибернетика и программирование. – 2016. – № 4. – С. 44–57.

23. Тарасов, Е. С. Методологические основы принятия решений с использованием автоматизации неформальных процедур / Е. С. Тарасов, В. С. Симанков, М. М. Путято // Естественные и технические науки. – 2010. – № 4 (49). – С. 292–297.

24. Тюкин, Ю. А. Адаптация в нелинейных динамических системах / Ю. А. Тюкин, В. А. Терехов. – Санкт-Петербург : ЛКИ. – 2008. – С. 384.

25. Учаев, Д. Ю. Анализ и управление рисками, связанными с информационным обеспечением человеко-машинных АСУ технологическими процессами в реальном времени / Д. Ю. Учаев, Ю. М. Брумштейн, И. М. Ажмухадедов, О. М. Князева, И. А. Дюдиков // Прикаспийский журнал: управление и высокие технологии. – 2016. – № 2. – С. 82–95.

26. Харченко, В.С. Оценка надежности информационно-управляющих систем на основе иерархических fmeicia-таблиц и марковских цепей: модели, методика и информационная технология / В. С. Харченко, И. Э. Комари, А. В. Горбенко // Научные ведомости белгородского государственного университета. Серия: экономика. информатика. – 2011. – № 19 (114). –С. 162–171.

27. Юрков, Н. К. К проблеме моделирования риска отказа электронной аппаратуры длительного функционирования / Н. К. Юрков, И. И. Кочегаров, Д. Л. Петрянин // Прикаспийский журнал: управление и высокие технологии. – 2015. – № 4. – С. 220–231.

References

1. Abzalov, A. V., Zhedunov, R. R. Metodika analiza predavariynykh situatsiy na tekhnologicheskikh obektakh upravleniya [Methodology of analysis of pre emergency situations at technological control facilities]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2013, no. 4, pp. 50–59.

2. Al-Busaidi, S. S. S., Voyakina, Yu. N., Ponomarev, S. V. K voprosu o podderzhke protsessa prinyatiya resheniya ob uluchshenii deyatelnosti v ispytatelnoy laboratorii [To the question of supporting the decision-making process on performance improvement in the testing laboratory]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2021, no. 1, pp. 27–45.

3. Artemenkov, S. L., Alkhimov, V. I., Baranov, S. N., Belyaeva, O. B., Dumin, P. N., Kornienko, P. A., Kuravskiy, L. S. (ed.), Malykh, S. B., Margolis, A. A., Marmalyuk, P. A., Panfilova, A. S., Popkov, S. I., Yurev, G. A. *Markovskie modeli v zadachakh diagnostiki i prognozirovaniya* [Markov models in diagnostic and forecasting tasks]. Moscow, Publishing House of the Moscow State Psychological and Pedagogical University, 2017. 203 p.

4. Bachmanov, D. A., Ocheredko, A. R., Putyato, M. M., Makaryan, A. S. Issledovanie voprosov sovershenstvovaniya sistem zashchity ot DDoS-atak na osnove kompleksnogo analiza sovremennykh mekhanizmov protivodeystviya [Research of the issues of improvement of protection systems against DDoS-attacks based on the comprehensive analysis of modern interaction mechanisms]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2021, no. 1, pp. 63–74. DOI: 10.21672/2074-1707.2021.53.1.063-074.

5. Bugaychenko, P. Yu., Danilov, A. I., Zubachev, A. M. Dinamicheskaya model otsenivaniya kachestva podgotovki i primeneniya slozhnoy tekhnicheskoy sistemy [Dynamic model for assessing the quality of preparation and application of a complex technical system]. *Trudy voenno-kosmicheskoy akademii im. A.F. Mozhayskogo* [Proceedings of the Military Space Academy named after A.F. Mozhaisky], 2018, no. 664, pp. 20–26.

6. Bunkin, N. F. *Stokhasticheskie sistemy v fizike i tekhnike* [Stochastic systems in physics and engineering]. Moscow, Publishing House of Bauman Moscow State Technical University, 2011. 366 p.

7. Volodin, I. V., Putyato, M. M., Makaryan, A. S., Evglevskiy, V. Yu. Klassifikatsiya mekhanizmov atak I issledovanie metodov zashchity sistem s ispolzovaniem algoritmov mashinnogo obucheniya i iskusstvennogo intellekta [Classification of attack mechanisms and investigation of system protection methods using machine learning and artificial intelligence algorithms]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2021, no. 2, pp. 91–98. DOI: 10.21672/2074-1707.2021.53.1.090-098.

8. Danilov, A. I., Zubachev, A. M., Danilov, A. A. Metodika chislennogo analiza effektivnosti podgotovki i primeneniya slozhnoy tekhnicheskoy sistemy [Methodology of numerical analysis of the effectiveness of preparation and application of a complex technical system]. *Izvestiya Tulskogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [Proceedings of Tula State University. Technical sciences], 2018, no. 8, pp. 186–199.

9. Iradzh, Elyasi Komari, Gorbenko, A. V. Analiz zadach razrabotki i reinzhiniringa kompyuternykh setey dlya kriticheskikh prilozheniy [Analysis of computer network development and reengineering tasks for critical applications]. *Radyuelektront i kompyutersh sistemi* [Radyuelectron i Computer Systems], 2006, no. 7 (19), pp. 32–35.

10. Klyatskin, V. I. *Ocherki po dinamike stokhasticheskikh system* [Essays on the dynamics of stochastic systems]. Moscow, Krassand Publ., 2012. 448 p.

11. Kochnev, S. V., Lapsar, A. P. Parametricheskiy podkhod k otsenke sostoyaniya slozhnykh tekhnicheskikh obektov na osnove evolyutsionnykh uravneniy [Parametric approach to assessing the state of complex technical objects based on evolutionary equations]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2017, no. 2, pp. 10–18.

12. Kubarev, A. V., Lapsar, A. P., Nazaryan, S. A. Parametricheskoe modelirovanie sostoyaniya obektov kriticheskoy infrastruktury v usloviyakh destruktivnogo vozdeystviya [Parametric modeling of the state of critical infrastructure objects in conditions of destructive impact]. *Voprosy kiberbezopasnosti* [Cybersecurity Issues], 2021, no. 3, pp. 58–67. DOI: 10.21681/2311-3456-2021-3-58-67.

13. Kubarev, A. V., Lapsar, A. P., Fedorova, Ya. V. Povyshenie bezopasnosti ekspluatatsii znachimykh obektov kriticheskoy infrastruktury s ispolzovaniem parametricheskikh modeley evolyutsii [Improving the safety of operation of significant critical infrastructure facilities using parametric models of evolution]. *Voprosy kiberbezopasnosti* [Cybersecurity Issues], 2020, no. 1, pp. 8–17. DOI: 10.21681/2311-3456-2020-01-08-17.

14. Kureychik, V. M. Osobennosti postroeniya sistem podderzhki prinyatiya resheniy [Features of building decision support systems]. *Izvestiya Yuzhnogo federalnogo universiteta. Tekhnicheskie nauki* [Bulletin of the Southern Federal University. Technical Sciences], 2012, Thematic issue, pp. 92–98.

15. Maksimenko, V. N., Yasyuk, E. V. Osnovnye podkhody k analizu i otsenke riskov informatsionnoy bezopasnosti [The main approaches to the analysis and assessment of information security risks]. *Ekonomika i kachestvo sistem svyazi* [Economics and Quality of Communication Systems], 2017, no. 2, pp. 42–48.

16. Mikheev, M. Yu., Prokofev, O. V., Savochkin, A. E., Linkova, M. A. Matematicheskie i informatsionno-strukturnye modeli prognozirovaniya sostoyaniya tekhnicheski slozhnykh obektov [Mathematical and information-structural prediction models state of technically complex objects]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 4, pp. 232–249.

17. Nikolaev, A. B., Yakunin, P. S., Krasnov, Yu. A. Avtomatizatsiya podderzhki upravlencheskikh resheniy pri organizatsii naukoemkogo proizvodstva na osnove gibkoy obratnoy svyazi [Automation of management decision support in the organization of knowledge-intensive production based on flexible feedback]. *Moskovskiy avtomobilno-dorozhnyy gosudarstvennyy tekhnicheskiy universitet.* [Moscow Automobile and Highway State Technical University], 2013, no. 4, pp. 3–13.

18. Oksendal, B. *Stokhasticheskie differentsialnye uravneniya. Vvedenie v teoriyu i prilozheniya* [Stochastic differential equations. Introduction to theory and applications]. Moscow, Mir Publ. ; AST Publ., 2003. 408 p.

19. Ostreykovskiy, V. A., Salnikov, N. L. *Veroyatnostnoe prognozirovanie rabotosposobnosti elementov YaEU* [Probabilistic prediction of the operability of elements YAU]. Moscow, Energoatomizdat Publ., 1990. 416 p.

20. Pugachev, V.S., Sinitsyn, I.N. *Stokhasticheskie differentsialnye sistemy* [Stochastic differential systems]. Moscow, Nauka Publ., 2000. 1000 p.

21. Savochkin, A. E. Algoritmizatsiya raboty sistemy monitoringa i kontrolya dlya resheniya zadach identifikatsii stepeni povrezhdeniya tekhnicheski slozhnykh obektov [Algorithmization of system operation of monitoring and control for solution of identification tasks of damage level of technically complex objects]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2014, no. 2, pp. 23–35.

22. Sibiryakov, M. A., Vasyaeva, E. S. Modifikatsiya i modelirovanie algoritmov obrabotki dannykh v keshpamyati sistem khraneniya dannykh [Modification and modeling of data processing algorithms in the cache memory of data storage systems]. *Kibernetika i programmirovanie* [Cybernetics and programming], 2016, no. 4, pp. 44–57. DOI: 10.7256/2306-4196.2016.4.18058.

23. Tarasov, E. S., Simankov, V. S., Putyato, M. M. Metodologicheskie osnovy prinyatiya resheniy s ispolzovaniem avtomatizatsii neformalnykh protsedur [Methodological foundations for decision-making using the automation of informal procedures]. *Estestvennye i tekhnicheskie nauki* [Natural and Technical Sciences], 2010, no. 4 (49), pp. 292–297.

24. Tyukin, Yu. A., Terekhov, V. A. *Adaptatsiya v nelineynykh dinamicheskikh sistemakh* [Adaptation in nonlinear dynamical systems]. Saint Petersburg, LKI Publ., 2008. 384 p.

25. Uchaev, D. Yu., Brumshteyn, Yu. M., Azhmukhadedov, I. M., Knyazeva, O. M., Dyudikov, I. A. Analiz i upravlenie riskami, svyazannymi s informatsionnym obespecheniem cheloveko-mashinnykh ASU tekhnologicheskimi protsessami v realnom vremeni [Analysis and management of risks, concerned with information support of human-machine automated control systems, operating in real time]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2016, no. 2, pp. 82–95.

26. Kharchenko, V. S., Komari, I. E., Gorbenko, A. V. Otsenka nadezhnosti informatsionno-upravlyayushchikh sistem na osnove ierarkhicheskikh fmeicia-tablits i markovskikh tsepey: modeli, metodika i informatsionnaya tekhnologiya [Reliability assessment of information and control systems based on hierarchical fmeicia-tables and Markov chains: models, methodology and information technology]. *Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta. Seriya: ekonomika. Informatika* [Scientific Bulletin of Belgorod State University. Series: Economics. Informatics.], 2011, no. 19 (114), pp. 162–171.

27.Yurkov, N. K., Kochegarov, I. I., Petryanin, D. L. K probleme modelirovaniya riska otkaza elektronnoy apparatury dlitelnogo funktsionirovaniya [Model for risk estimation of failure of electronic means, intended for long-term performance]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 4, pp. 220–231.

DOI 10.54398/2074-1707_2022_1_106
УДК 681.3.06

АНАЛИЗ ВОЗМОЖНОСТЕЙ МОНИТОРИНГА ЛОКАЛЬНОЙ ДОМЕННОЙ СЕТИ КЛАССА С СРЕДСТВАМИ MICROSOFT POWERSHELL

Статья поступила в редакцию 18.01.2022, в окончательном варианте – 17.02.2022.

Григорьев Роман Игоревич, Уральский технический институт связи и информатики (филиал) ФГБОУ ВО «Сибирский государственный университет телекоммуникаций и информатики», 620109, Российская Федерация, г. Екатеринбург, ул. Репина, 15,
студент, ORCID: 0000-0002-2835-3807, e-mail: c.terner@mail.ru
Осипова Ирина Александровна, Уральский технический институт связи и информатики (филиал) ФГБОУ ВО «Сибирский государственный университет телекоммуникаций и информатики», 620109, Российская Федерация, г. Екатеринбург, ул. Репина, 15,
кандидат технических наук, доцент, ORCID: 0000-0002-3823-1141, e-mail: minesur@mail.ru

В статье рассмотрены этапы разработки пользовательского модуля Powershell для мониторинга сети класса С. Проанализированы текущие проблемы в локальной сети библиотеки им. В.Г. Белинского. Сформированы технические задачи системного администратора по их устранению. Указаны особенности средств автоматизации Powershell. Проанализированы предварительные действия для развертывания Powershell. Предложена концепция процесса сетевого мониторинга. На основе задач разработаны некоторые функции модуля Powershell и рассмотрен принцип проектирования других. Протестированы функции обнаружения устройств в сети и формирования отчета о рабочих станциях под управлением Windows в формате HTML. Рассмотрен репозиторий пользовательских модулей PowershellGallery. Показана структура работы модуля и порядок действий при его использовании.

Ключевые слова: Powershell, средство автоматизации, технология WMI, локальная сеть, отчет о компьютере, формат отчета HTML, модуль Powershell, функции Powershell, сбор информации о компьютерах, обнаружение устройств в сети, PowershellGallery, система мониторинга

ANALYSIS OF FEATURES OF MONITORING OF LOCAL DOMAIN NETWOK CLASS C BY USING THE TOOL MICROSOFT POWERSHELL

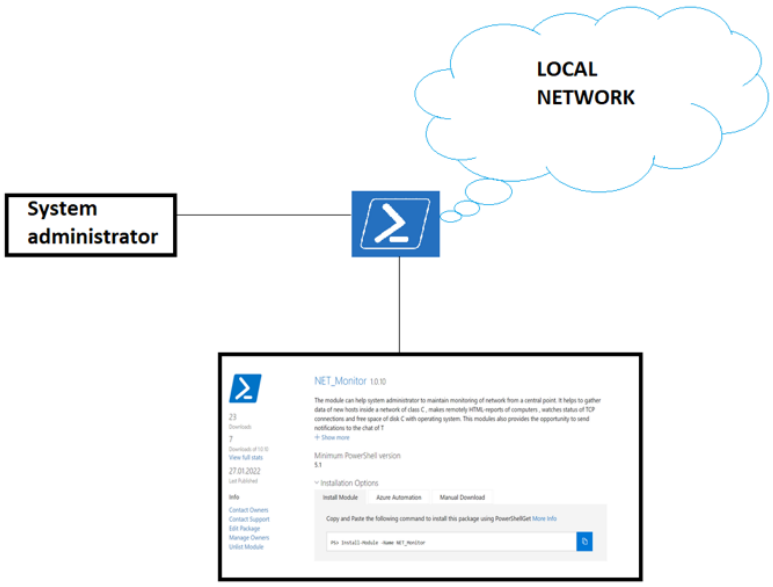
The article was received by the editorial board on 18.01.2022, in the final version – 17.02.2022.

Grigoriev Roman I., Ural Technical University of Connection and Informatics, 15 Repin St., Yekaterinburg, 620109, Russian Federation,
student, ORCID: 0000-0002-2835-3807 e-mail: c.terner@mail.ru
Osipova Irina A., The Ural Technical University of Connection and Informatics, 15 Repin St., Yekaterinburg, 620109, Russian Federation,
Cand. Sci. (Engineering), Assistant Professor, ORCID: 0000-0002-3823-1141, e-mail: minesur @mail.ru

The article shows stages of development of the custom Powershell module which is designed for monitoring networks. It analyzed current problems of local network of SOUNB named after V.G. Belinskiy. The technical tasks for system administrator were also formedto fix current problems. The article considers features of automation tool Powershell. It analyzed preparatory actions for deployment of Powershell. The conception of process of monitoring of network were offered to consideration. Depends on tasks some functions of the module were constructed, and schematics of others were considered. The functions of detection of other devices and of making HTML reports of working stations were tested. The article considers the repository of custom modules PowershellGallery. The structure of work of the module and actions of its deployment were demonstrated.

Keywords: Powershell, automation tool, WMI technology, local network, report of computer, format of report HTML, Powershell module, Powershell functions, gathering data of computers, detection of devices, PowershellGallery, system of monitoring

Графическая аннотация (Graphical annotation)



Введение. В данной работе исследуется возможность создания персонализированных инструментов мониторинга локальной сети класса С для небольших предприятий без использования стороннего программного обеспечения и комплексных систем мониторинга. Улучшенные функции Microsoft Powershell с возможностями командлетов группируются в отдельный модуль, который загружается в центральный репозиторий Powershell Gallery. Системный администратор сможет самостоятельно определять, какой функционал ему необходим для поддержания сетевой инфраструктуры. В дальнейшем, при помощи одной команды Install-Module, такой модуль можно установить на любой компьютер в локальной сети и поддерживать централизованное управление всеми хостами из одной точки сети. Потенциал применения таких функций позволяет расширить функционал библиотек Microsoft Powershell и, при использовании платных графических средств Powershell, даже позволит реализовать собственную систему мониторинга, скомпилированную в exe-файл при помощи модуля Powershell ps2exe.

Процесс поддержания работоспособности локальной вычислительной сети на предприятии является основной ответственностью системного администратора. Для того чтобы эффективно устранять возникающие неполадки, специалист должен периодически проводить мониторинг сети с целью обновления списка устройств в сети и информации о каждом из них. Наиболее используемыми форматами для обработки информации являются текстовые и табличные, в то время как для взаимодействия с человеком предпочтительно использование HTML-формата гиперссылок.

На настоящий момент времени существует множество систем мониторинга – платных и с открытым кодом. В работе Лукаса Макуры “Multi-riteria Analysis and Prediction of Network Incidents Using Monitoring System” и Эмануэля Згардея “Improving IT Infrastructure Management Using Nagios Open Source Package” были рассмотрены наиболее оптимизированные решения мониторинга с открытым исходным кодом – Nagios, Zabbix [1, 2]. Среди платных продуктов можно выделить Network Olympus, Paessler PRTG Network Monitor, 10-Страйк Мониторинг и др.

При рассмотрении тенденций обеих этих групп у каждой из них можно выделить ряд характерных признаков, которые показывают определенные достоинства и недостатки. Системы мониторинга с открытыми кодом бесплатны, имеют объемное сообщество и полную документацию. К их недостаткам можно отнести достаточно долгий промежуток между обновлениями или исправлениями багов, их многокомпонентную структуру, медленное развертывание, недоступность выполнения некоторых функций мониторинга без дополнительных действий или навыков, высокое потребление ресурсов и работа с агентами. Платные решения скомпилированы в одну программу или набор готовых к использованию модулей, способны к быстрому развертыванию и использованию, содержат полный набор необходимых функций. Среди недостатков подобных систем выделяются система ежемесячной/ежегодной подписок, возможность работать лишь с GUI и зависимость от внешней службы поддержки продукта.

Не всем ресурсам локальной сети требуется мониторинг. Его необходимость возникает из поставленных задач, требующих лаконичного подхода. В сетях класса С при лимите до 1000 рабочих

машин количество критически важных компонентов сети крайне мало и не требует использования всех предлагаемых сторонними продуктами функций. Однако такие системы также должны быть способны выполнять базовый набор задач, которые использует в своей работе системный администратор. К ним можно отнести обнаружение устройств в сети, проектирование отчетов об аппаратных комплектующих сетевых устройств и др.

В условиях ограниченности ресурсов кластера, недоступности некоторых функций мониторинга и времени развертывания сторонних продуктов появляется необходимость в создании дополнительных инструментов мониторинга, позволяющих автоматизировать определенные задачи системного администратора. Разработка некоторых из них и методы их внедрения будут рассмотрены на примере локальной сети предприятия СОУНБ им. В.Г. Белинского.

Предлагаемая методика. Создание подобных инструментов в компьютерах под управлением Windows стало возможным благодаря средству автоматизации Powershell от Microsoft. Оно позволяет инкапсулировать данные о системе в объекты и использовать их в дальнейшем при обработке запросов. Powershell – полноценный язык программирования, позволяющий использовать классы других языков программирования и специальные командлеты для доступа к ресурсам операционной системы.

Командлеты – это улучшенный тип функций, которые позволяют определенным образом передавать им аттрибуты и передавать их вывод другим функциям с помощью конвеера. Конвеер Powershell – это уникальная особенность Powershell, которая делает возможным передачу вывода одной команды на вход следующей. Она позволяет выстраивать интересные конструкции и разбивать сложные задачи на несколько этапов.

Средство автоматизации Powershell входит в состав компонентов Windows, однако требует несколько подготовительных этапов перед развертыванием. Среди всех языков программирования количество постов на StackOverFlow занимает лидирующее место в разделе безопасности [3, 4]. Прежде всего необходимо помнить, что политика использования сценариев Powershell в Windows по умолчанию запрещена – Restricted. Для того чтобы выполнить её развертывание, необходимо использовать административные шаблоны ActiveDirectory, либо воспользоваться следующим командлетом: Set-Execution Policy – Execution Policy Remote Signed (Unrestricted).

Рекомендуется оставлять политику Remote Signed, позволяющую использовать написанные вами лично сценарии без запроса сертификатов подтверждения. Политика выполнения скриптов Unrestricted позволяет использовать любые файлы сценариев Powershell, что является опасной практикой.

Для удаленного доступа к другим компьютерам с помощью технологии сеансового уровня, реализованного в Powershell, необходимо добавить на локальном компьютере имена всех удаленных компьютеров, к которым будет выполняться подключение, в список доверенных хостов с помощью определенного командлета с атрибутом:

Set-Item WSMan:\localhost\Client\TrustedHosts – Value (hostname).

Помимо этого, на каждом из удаленных компьютеров должна быть запущена служба WMIили IIS для обработки подобных подключений.

Кроме того, необходимо открыть порт брендмауэра 5985 или выполнить в командной строке команду “winrmquickconfig”. Она создает исключение фаервола для текущего пользователя.

Еще одним шагом при подготовке является синхронизация установленных версий Powershell. В комплекте Windows 10, начиная с 19 версии, уже встроен Powershell 5.1, в то время как на Windows 7 может стоять лишь Powershell 2.0. Реализация технологии WMI для извлечения информации о компьютере появилась лишь начиная с третьей версии, как и некоторые способы вызова циклов. В версиях 1.0–5.1 Powershell считается привязанным к ОС Windows, однако шестая версия Powershell Core является уже кроссплатформенной. Язык Powershell основан на NetFramework и использует классы языка C#. С выходом NetCoreC# стал кросс-платформенным, что перестало привязывать Powershell только к одной операционной системе. Начиная с Powershell 6.0, он стал доступен на MacOs, Windows, Linux [5–7].

После проделанных действий можно писать собственные сценарии под конкретно поставленный набор задач системного администратора. Поскольку они будут выполняться многократно и автоматически, их можно реализовать в виде функций, вызываемых по требованию или необходимости. Ранее уже было сказано, что в языке программирования Powershell есть два типа функций – обычные и улучшенные. Начиная с версии 3.0 языка Powershell, обычные функции перестали использоваться, так как их функционал уже содержится в улучшенных функциях, включающих в себя возможности командлетов.

Для определения необходимых функций нужно сформировать концепцию мониторинга сети и соотнести с техническим заданием предприятия.

Основные этапы сетевого мониторинга представлены на рисунке 1, они будут циклически выполняться для каждого устройства.

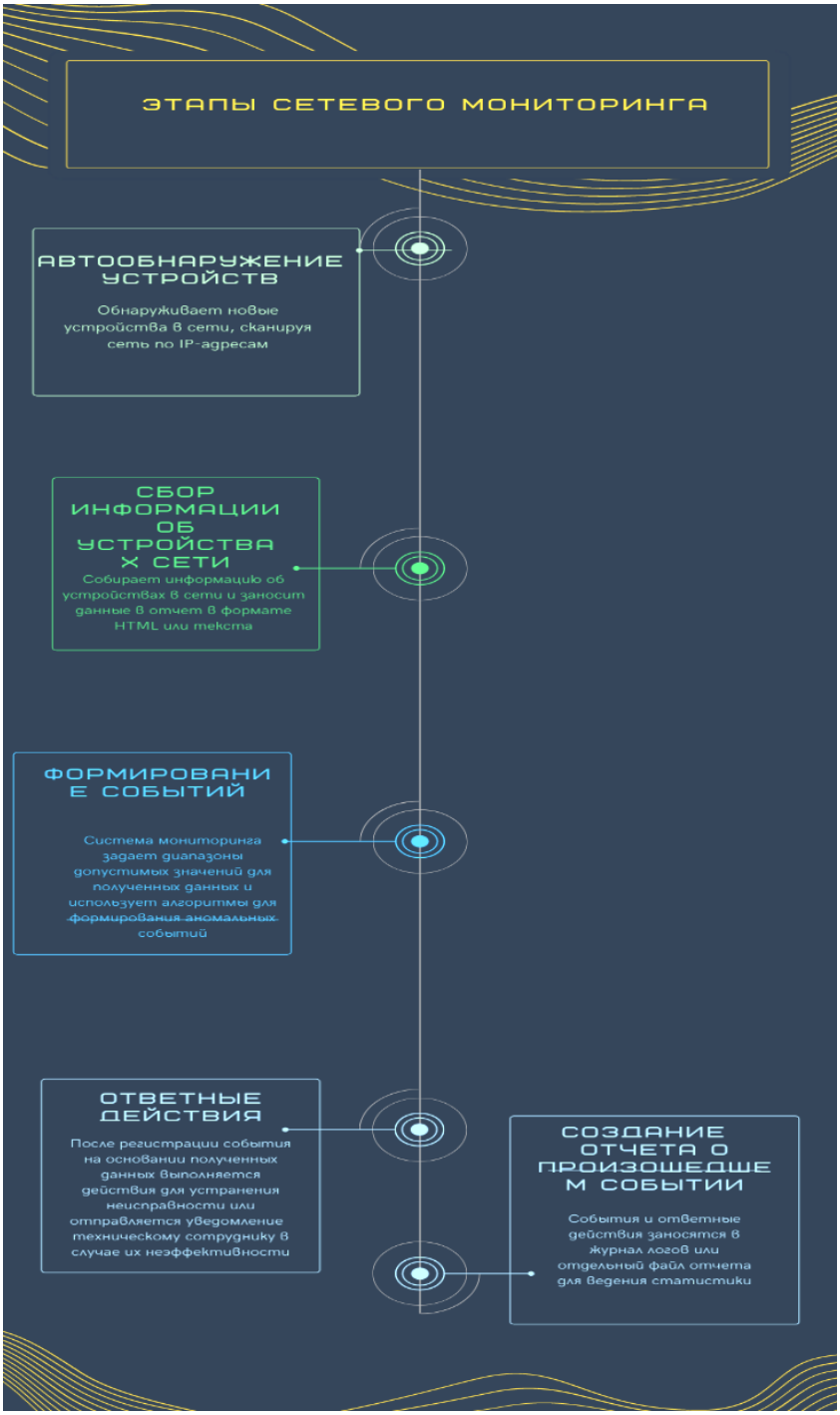


Рисунок 1 – Основные этапы сетевого мониторинга

После осуществления развертывания Powershell и выполнения предварительных действий – установки политики запуска скриптов и заведения списка доверенных хостов – администратор пишет код функций, ориентированный на выполнение конкретных задач на разных этапах:

1. Обнаружение устройств и их экспорт в файл таблицы хранения паролей.
2. Создание отчетов о найденных устройствах в формате HTML.
3. Формирование событий об активных соединениях на TCP-портах удаленного доступа – SSH и RDP.

4. Формирование событий об изменении скорости текущего соединения с внешними доменами.
5. Отправка уведомлений о произошедших событиях в групповой чат Telegram.
6. Добавление записи о произошедшем событии в текстовый файл логов.

7.1. **Функция обнаружения устройств и их экспорт в файл таблицы хранения паролей.** В библиотеке реализованы пять сетей класса С – читательская, для сотрудников, серверная, видеонаблюдения и IP-телефонии. Сеть расположена в домене, но DHCP-роль не установлена. Все данные о пользователях и их паролях хранятся на сетевом ресурсе в таблице. На настоящий момент в библиотеке происходит реорганизация нескольких отделов. Возникла необходимость в автоматизированном обнаружении сетевых устройств и экспорта полученных хостов в табличный файл, в котором будет также храниться информация о пользователях и паролях. Пароли переносятся вручную в целях безопасности, а остальные данные – IP, DNS-имя, MAC, информация о процессоре, имя используемого пользователя – запрашиваются с каждого хоста сети с помощью технологии WMI в цикле for. Диапазон переменной счетчика i задается атрибутами функции – начало и конец диапазона сканирования. Технология WMI представляет собой репозиторий, место в котором разделено на различные адресные пространства для каждого класса устройств. Обратившись к пространству класса, с удаленного компьютера можно получить доступ к атрибутам конкретного устройства. Например, для класса win32_physicalmemory можно запросить любые атрибуты оперативной памяти – частота, объем, используемый слот на материнской плате компьютера. В качестве табличного файла можно использовать Excel или CSV. Последний был выбран из-за поддержки его встроенным модулем Powershell.

7.2. **Функция создания отчетов о найденных устройствах в формате HTML.** Следующая функция запрашивает уже более детальный отчет, используя ту же технологию, но экспортируя заранее отобранные атрибуты во внешний файл формата HTML. До настоящего момента отчеты об аппаратных компонентах компьютера снимались программами Everest и AIDA64 в момент физического нахождения рядом с устройством, чаще всего это происходило в момент диагностики устройства. Они хранились локально и запрашивались при необходимости для отслеживания замены компонентов компьютеров. Эти программы поддерживали и HTML-формат отчетов, но содержали в себе достаточно много лишней информации и имели объем до нескольких десятков мегабайт.

Формат гиперссылок считается более привычным для человеческого глаза и может быть использован на веб-сервере IIS для предоставления статистики. Процесс осуществляется за счет командлета ConvertTo-HTML. В строку вывода этой команды можно добавить html-теги, разбив текст на отдельные блоки. К ним можно применить стили CSS, которые могут быть включены непосредственно в скрипт Powershell или прикрепляться с помощью внешнего файла с расширением CSS. Кроме того, такие отчеты не содержали информацию о производительности устройств в реальном времени – работе их процессов и служб. Сотрудниками отдела технического обеспечения СОУНБ им. В.Г. Белинского был выведен ряд параметров, которые требовалось включить в содержание отчета о рабочих станциях. Заключительный вариант включал в себя три раздела – информация о ключевых аппаратных компонентах, производительность на уровне служб и процессов на момент снятия отчета, а также подключенные к компьютеру принтеры.

Отчеты снимались по пять минут, и информацию из них было крайне трудно извлечь. Это представлено на рисунке 2.

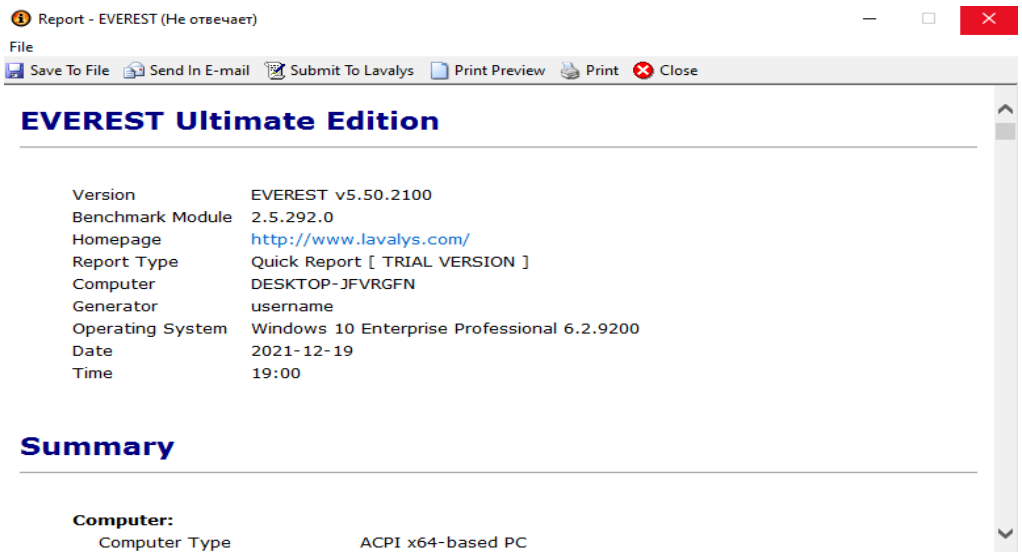


Рисунок 2 – Скриншот отчета информации

7.3. **Функция формирования событий об активных соединениях на TCP-портах удаленного доступа – SSH и RDP**. Эти две функции включают в себя первые два этапа описанной концепции. В качестве формирования событий могут быть использованы параметры статуса работы служб, состояние активных соединений на определенных TCP портах, скорости соединения с внешним доменом. Концепция этой функции построена на открытии удаленной сессии с другим компьютером, его периодическим опросом в бесконечном цикле через определенные интервалы времени.

7.4. **Функция формирования событий об изменении скорости текущего соединения с внешними доменами.** В библиотеке подключены два провайдера – Erlang и Planeta. Скорость тарифов на каждом из них – 30 и 100 Мбит/соответственно. При осуществлении перехода между ними будет сформировано событие, при котором скорость соединения сервера достигнет определенного значения. Для реализации этой функции можно использовать классы C# – Webclient или System.Net.Http.HttpClient из библиотеки SystemNet.

7.5. **Функция отправки уведомлений о произошедших событиях в групповой чат Telegram.** Для быстрого реагирования системному администратору необходима функция уведомлений, которая будет отправлять сообщения не на почту, а на почтовый мессенджер Telegram. Общение между техническими специалистами библиотеки осуществляется с помощью группового чата, потребуется создать бота в телеграме через специальный чат, далее использовать его токен – уникальный идентификатор, набор вызовов API для принятия и отправки сообщений.

7.6. **Функция добавления записи о произошедшем событии в текстовый файл логов.** Необходима функция последнего этапа – запись произошедшего события в логи для составления статистики. Она выполняется с помощью командлета Set-Content с атрибутом добавления строки в файл.

Описание структуры модуля. Пользовательский модуль Powershell – это файл с расширением psm1, в котором расположен листинг всех включенных в него функций. Он хранится в одном из трех стандартных расположений на локальном компьютере. Начиная с третьей версии Powershell достаточно использовать функцию модуля, чтобы все его функции из физического расположения загрузились в сеанс консоли автоматически.

После создания файла модуля можно дополнительно к нему создать файл манифеста модуля с раширением psd1. В нем хранится дополнительная информация о созданном модуле – поддерживаемая версия Powershell, информация об авторе и краткое описание его работы.

Реализация работы первых этапов модуля. После развертывания системному администратору нужно войти в систему под учетной записью с соответствующими правами; сравнить установленную версию Powershell с версией, указанной в манифесте модуля; открыть консоль Powershell или среду ISE, запустив с правами администратора; один раз установить модуль с помощью командлета Install-Module и подтвердить действие в консоли; вызывать и использовать функции этого модуля при составлении других сценариев. При повторном использовании Powershell на этом компьютере устанавливать модуль уже не требуется – достаточно вызвать любую из его функций. Скриншоты работы первых двух функций приведены на рисунках 3–6.

Итоговым этапом будет регистрация модуля в общедоступном репозитории PowershellGallery. Это необходимо для того, чтобы иметь возможность устанавливать его на разных компьютерах с помощью одного командлета Install-Module.

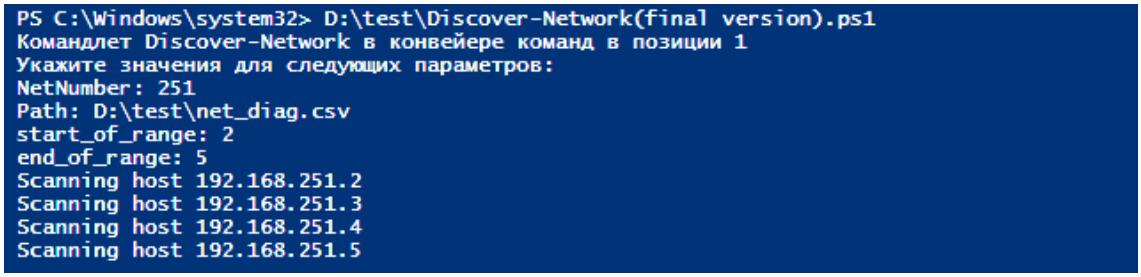
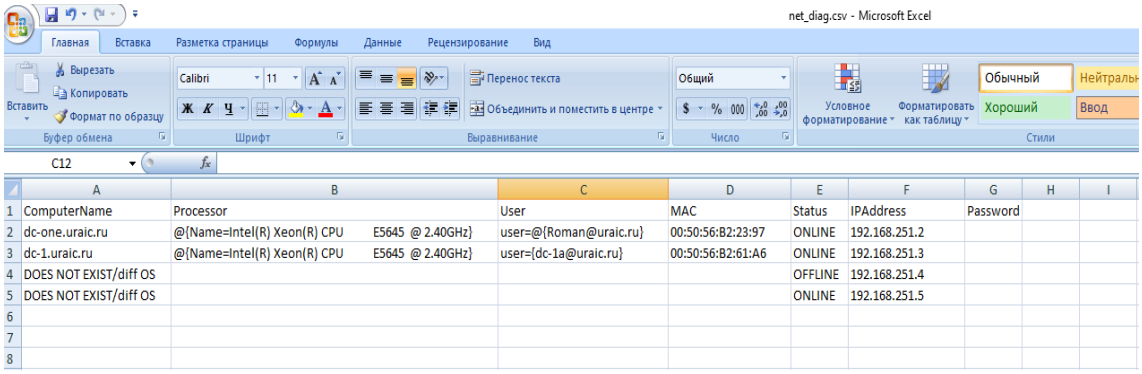


Рисунок 3 – Скриншот выполнения функции обнаружения устройств и их экспорт в файл таблицы хранения паролей



net_diag.csv - Microsoft Excel

	A	B	C	D	E	F	G	H	I
1	ComputerName	Processor	User	MAC	Status	IPAddress	Password		
2	dc-one.uraic.ru	@{Name=Intel(R) Xeon(R) CPU E5645 @ 2.40GHz}	user=@{Roman@uraic.ru}	00:50:56:B2:23:97	ONLINE	192.168.251.2			
3	dc-1.uraic.ru	@{Name=Intel(R) Xeon(R) CPU E5645 @ 2.40GHz}	user={dc-1a@uraic.ru}	00:50:56:B2:61:A6	ONLINE	192.168.251.3			
4	DOES NOT EXIST/diff OS				OFFLINE	192.168.251.4			
5	DOES NOT EXIST/diff OS				ONLINE	192.168.251.5			
6									
7									
8									

Рисунок 4 – Скриншот результата функции обнаружения устройств и их экспорт в файл таблицы хранения паролей

```
PS C:\WINDOWS\system32> hostname
DESKTOP-JFVRGFN

PS C:\WINDOWS\system32> D:\Учеба\ВКР(Powershell)\ИТОГОВОЕ\CreateHTMLreport-Computer.ps1
cmdlet CreateHTMLreport-Computer at command pipeline position 1
Supply values for the following parameters:
ComputerName: DESKTOP-JFVRGFN
Path: D:\Powershell
Making report for DESKTOP-JFVRGFN ...

PS C:\WINDOWS\system32>
```

Рисунок 5 – Скриншот выполнения функции снятия отчета компьютера в формате HTML

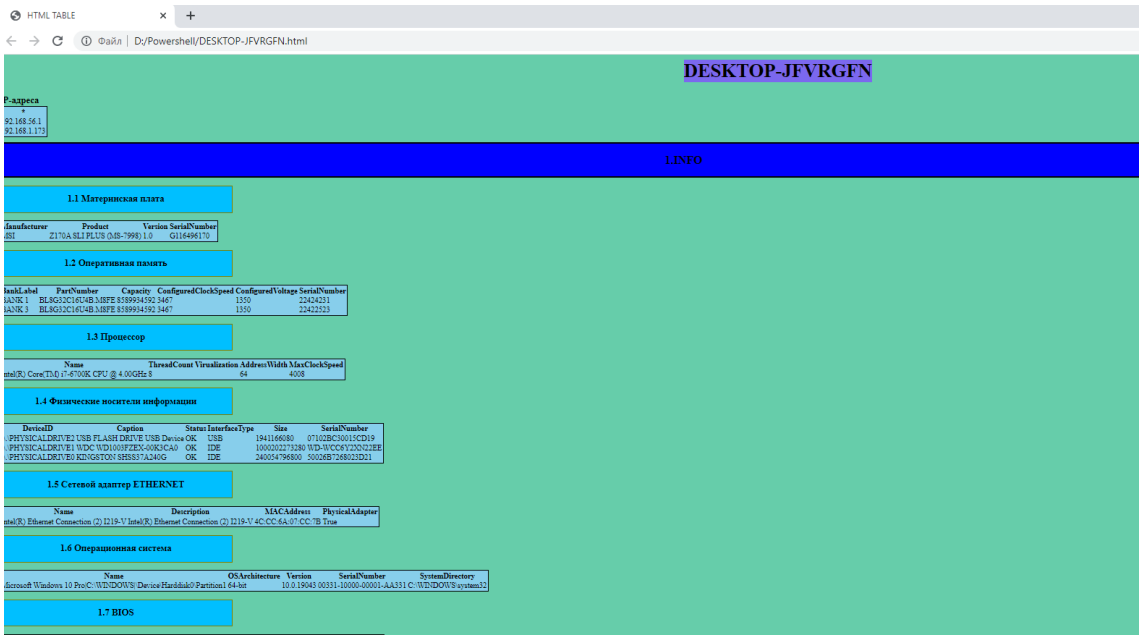


Рисунок 6 – Скриншот результата функции снятия отчета компьютера в формате HTML

Таким образом, этот модуль станет набором программных инструментов для системных администраторов с открытым исходным кодом и дополнит функциональные возможности средства автоматизации Powershell.

Заключение. В работе предложены этапы разработки пользовательского модуля Powershell для мониторинга сети класса С. Проанализированы текущие проблемы в локальной сети библиотеки им. В.Г. Белинского, а также предварительные действия для развертывания Powershell.

Решения задач разработки некоторых функций модуля Powershell и рассмотрения принципа проектирования других осложняются тестированием функции обнаружения устройств в сети и формирования отчета о рабочих станциях под управлением Windows в формате HTML.

Помимо этого, рассмотрен репозиторий пользовательских модулей PowershellGallery. Показана структура работы модуля и порядок действий при его использовании.

Библиографический список

1. Lukas, Macura, Miroslav, Voznak. Multi-criteria analysis and prediction of network incidents using monitoring system / Lukas Macura, Miroslav Voznak // Journal of Advanced Engineering and Computation. – 2019, October. – Vol. 1, № 1. – P. 31–34.
2. Emanuel, Zgârdea, Ladislau, Augustinov. Improving IT Infrastructure Management Using Nagios Open Source Package / Emanuel Zgârdea, Ladislau Augustinov // Analele Universităţii "Eftimie Murgu" Reşiţa: Fascicola I, Inginerie. – 2014, December. – P. 334–336.
3. Sudhakar, Sushil, Kumar. An emerging threat Fileless malware: a survey and research challenges / Sudhakar, Sushil Kumar // Cybersecurity. – 2020, February. – P. 1–4.
4. Croft, Roland. An empirical study of developers' discussions about security challenges of different programming languages / Croft Roland, Xie Yongzheng, Zahedi Mansooreh, Babar M. Ali, Treude // Empirical Software Engineering. – 2021, December. – P. 23–27.
5. The Power Shell Gallery. – Режимдоступа: https://docs.microsoft.com/en-us/powershell/scripting/gallery /overview?view=powershell-7.2, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 22.12.2021).
6. Бертрам, А. Powershell для сисадминов / А. Бертрам. – Санкт-Петербург : Издательский дом «Питер», 2021. – 416 с.
7. Олифер, В. Компьютерные сети. Принципы, технологии, протоколы / В. Олифер, Н. Олифер. – 5-е изд. – Санкт-Петербург : Издательский дом «Питер», 2016. – 992 с.

References

1. Lukas, Macura, Miroslav, Voznak. Multi-Criteria Analysis and Prediction of Network Incidents Using Monitoring System. *Journal of Advanced Engineering and Computation*, 2019, October, vol. 1, no. 1, pp. 31–34.
2. Emanuel, Zgârdea, Ladislau, Augustinov. Improving IT Infrastructure Management Using Nagios Open Source Package. *Analele Universităţii "Eftimie Murgu" Reşiţa: Fascicola I, Inginerie*, 2014, December, pp. 334–336.
3. Sudhakar, Sushil, Kumar. An emerging threat Fileless malware: a survey and research challenges. *Cybersecurity*, 2020, February, pp. 1–4.
4. Croft, Roland, Xie, Yongzheng, Zahedi, Mansooreh, Babar, M. Ali, Treude, Christoph. An empirical study of developers' discussions about security challenges of different programming languages. *Empirical Software Engineering*, 2021, December, pp. 23–27.
5. The Powershell Gallery. Available at: https://docs.microsoft.com/en-us/powershell/scripting/gallery /overview?view=powershell-7.2 (accessed 21.12.2021).
6. Bertram, A. *Powershell dlya sisadminov* [Powershell for sysadmins]. Saint-Petersburg, Publishing House "Piter", 2021. 416 p.
7. Olifer, V., Olifer, N. *Kompyuternye seti. Printsipy tekhnologii, protokoly* [Computer networks. Principles, technologies and protocols]. Saint-Petersburg, Publishing House "Piter", 2016. 992 p.

DOI 10.54398/2074-1707_2022_1_113
УДК 004.001

МЕТОД ЗАЩИТЫ СИСТЕМЫ МАШИННОГО ОБУЧЕНИЯ ОТ ВРЕДОНОСНЫХ ПРОГРАММ

Статья поступила в редакцию 01.02.2022, в окончательном варианте – 15.02.2022.

Петренко Вячеслав Иванович, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

кандидат технических наук, и. о. директора Института цифрового развития, заведующий кафедрой организации и технологии защиты информации, ORCID: 0000-0003-4293-7013, e-mail: vipetrenko@ncfu.ru

Тебуева Фариза Биляловна, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

доктор физико-математических наук, заведующая кафедрой компьютерной безопасности, ORCID: 0000-0002-7373-4692, e-mail: ftebueva@ncfu.ru

Анзоров Артур Русланович**,** Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

студент, ORCID: 0000-0002-1157-4021, e-mail: artanzrv@gmail.com

Стручков Игорь Владиславович, Северо-Кавказский федеральный университет, 355017, Российская Федерация, г. Ставрополь, ул. Пушкина, 1,

аспирант, ORCID: 0000-0001-8744-498X, e-mail: selentar@bk.ru

Статья посвящена проблеме защиты системы машинного обучения от вредоносного ПО. Проведен анализ возможных уязвимостей систем машинного обучения, приведена классификация наиболее опасных атак с описанием классов, включающих в себя способ воздействия и последствия от применения данных атак в системе машинного обучения. Для противодействия ряду атак предложен метод защиты системы машинного обучения от вредоносных программ на основе алгоритмов Neural-Cleanse и Jpeg-Compression.

Ключевые слова: машинное обучение, нейронные сети, информационная безопасность, Neural-Cleanse, Jpeg-Compression, атаки отравления, атаки искажения, атаки извлечения модели

A METHOD FOR PROTECTING A MACHINE LEARNING SYSTEM FROM MALICIOUS PROGRAMS

The article was received by the editorial board on 01.02.2022, in the final version – 12.02.2022.

Petrenko Vyacheslav I., North-Caucasian Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

Candidate Sci. (Engineering), Acting Director of the Institute for Digital Development, Head of the Department of Organization and Technology of Information Security, ORCID: 0000-0003-4293-7013, e-mail: vipetrenko@ncfu.ru

Tebueva Fariza B., North-Caucasian Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

Doct. Sci. of (Physics and Mathematics), Head of the Department of Computer Security, ORCID: 0000-0002-7373-4692, e-mail: ftebueva@ncfu.ru

Anzorov Artur R., North-Caucasian Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

student, ORCID: 0000-0002-1157-4021, e-mail: artanzrv@gmail.com

Struchkov Igor V., North-Caucasian Federal University, 1 Pushkin St., Stavropol, 355017, Russian Federation,

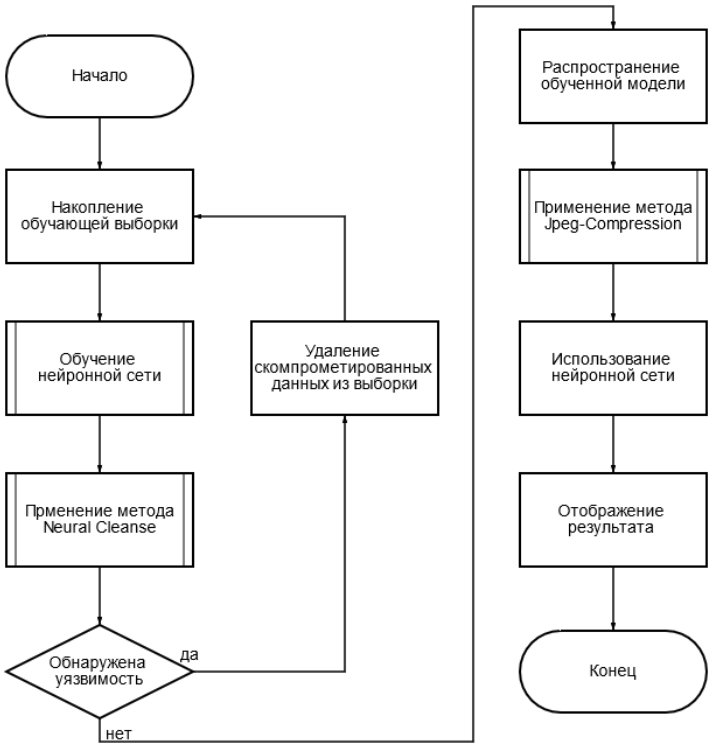
postgraduate student, ORCID: 0000-0001-8744-498X, e-mail: selenbar@bk.ru

The article is devoted to the problem of protecting a machine learning system from malware. An analysis of possible vulnerabilities of machine learning systems has been carried out, a classification of the most dangerous attacks with a description of classes, including the method of impact and consequences of using these attacks in a machine learning system, has been given. To counter a number of attacks, a method is proposed for protecting a machine learning system from malware based on the Neural-Cleanse and Jpeg-Compression algorithms.

Keywords: machine learning, neural networks, information security, Neural-Cleanse, Jpeg-Compression, poisoning attacks, evasion attacks, model extraction attacks

Graphical annotation (Графическая аннотация)

Метод защиты системы машинного обучения от вредоносных программ



Method for protecting a machine learning system from malware

Введение. За последние несколько лет машинное обучение стало заметным технологическим инструментом в нескольких прикладных областях, таких как компьютерное зрение, распознавание речи, понимание естественного языка, рекомендательные системы, поиск информации, компьютерные игры, медицинская диагностика, анализ рынка и т. д. Изучение и построение моделей с использованием обучающих данных дает злоумышленникам возможность атаковать алгоритмы машинного обучения, играя с функциями и границами решений модели. Злоумышленник может создавать вредоносные входные данные для атаки на производительность или эффективность алгоритма машинного обучения. Они могут обойти уже обученную систему, создав входные данные, которые используют систему для совершения существенных ошибок.

Поскольку машинное обучение становится важным инструментом в стратегически важных приложениях, таких как наблюдение за безопасностью и проверка биографических данных для решений о выдаче визы и т. д., важно понимать эти атаки и делать алгоритмы машинного обучения более устойчивыми к этим атакам. Если хакер еще не знает алгоритм, он сначала пытается изучить алгоритм и лежащую в его основе модель (например, логистическую регрессию, нейронную сеть, деревья решений и т. д.). Иногда хакер может быть заинтересован только в изучении модели, чтобы он мог построить свою собственную «копию» системы, используя изученную модель. Это может быть полезно, если приложение предлагается как услуга через API и с пользователей взимается плата за использование этих API.

Хакер может создать последовательность входных данных, а затем, наблюдая выходные данные системы, соответствующие этим входным данным, он может построить локальную модель, которая может быть очень близка к модели, используемой исходной системой. В зависимости от цены и условий лицензии на использование API, хакер может «украсть» модель за очень небольшую сумму денег. Из-за быстрого развития технологий атак на системы на основе ИИ существующие меры защиты могут со временем стать менее эффективными, но подходы к защите и их обоснование остаются. Кроме того, большинство представленных подходов исходят из академической среды и делают определенные допущения, которые необходимо учитывать при применении этих подходов на практике.

Атаки на системы машинного обучения. Уязвимости в системах машинного обучения могут возникать из-за ошибок при проектировании, из-за неизбежных алгоритмических ограничений или из-за сочетания обоих этих факторов [1]. В настоящее время во всем мире работает множество систем машинного обучения – от «прозрачных ящиков», имеющих полностью открытый код и обученных на общедоступных данных, до «черных», получающих входные данные по закрытым протоколам после обработки неустановленными функциями преобразования и передающих результаты через проприетарные API. Между двумя этими крайностями есть другие варианты, в том числе системы машинного обучения с открытым кодом, но проприетарными гиперпараметрами и обучающими данными, а также «черные ящики», действующие по принципу переноса обучения с «прозрачных ящиков» [2].

Атаки на системы, основанные на машинном обучении, можно разделить на 4 основные категории [3]:

1) атака искажения (evasion attack);
2) атака отравления (poisoning);
3) атака инверсии модели (model inversion);
4) атака извлечения модели (model extraction).

Атаки искажения являются наиболее распространенным типом атаки на систему машинного обучения. Вредоносные входные данные тщательно обрабатываются, чтобы избежать обнаружения, что, по сути, означает, что входные данные изменяются таким образом, чтобы алгоритм машинного обучения классифицировал их как безопасные, а не вредоносные, заставив сеть выдавать неверные ответы в определенных ситуациях. Использование враждебного пространства (рис. 1) для поиска враждебных (аномальных) экземпляров, которые приводят к неправильной классификации в классификаторе на основе машинного обучения, является основой атак искажения.

Атака отравления происходит, когда злоумышленник может внедрить неверные данные в тренировочный пул вашей модели и, следовательно, заставить его учиться чему-то, чего он не должен. Атаки отравления могут быть проведены в двух вариантах – те, которые нацелены на доступность модели машинного обучения, и те, которые нацелены на его целостность (также известные как атаки типа бэкдор). Целью злоумышленника является искажение выходных данных модели машинного обучения, добавив в систему зараженные данные, что приведет к ошибочной классификации. При отравлении злоумышленники пытаются повлиять на данные обучения, чтобы повлиять на результат обучения. Отравляющие атаки могут варьироваться от влияния на производительность алгоритма обучения до преднамеренного внесения в модель определенных смещений.

Рисунок 1 – Враждебное пространство как результат несовершенного представления тренировочных данных

Атака отравления происходит, когда злоумышленник может внедрить неверные данные в тренировочный пул вашей модели и, следовательно, заставить его учиться чему-то, чего он не должен. Атаки отравления могут быть проведены в двух вариантах – те, которые нацелены на доступность модели машинного обучения, и те, которые нацелены на его целостность (также известные как атаки типа бэкдор). Целью злоумышленника является искажение выходных данных модели машинного обучения, добавив в систему зараженные данные, что приведет к ошибочной классификации. При отравлении злоумышленники пытаются повлиять на данные обучения, чтобы повлиять на результат обучения. Отравляющие атаки могут варьироваться от влияния на производительность алгоритма обучения до преднамеренного внесения в модель определенных смещений.

Атаки инверсии данных – это атаки, при которых злоумышленники, имеющие возможность запрашивать адресную модель, пытаются вывести набор обучающих данных. Эти атаки нарушают конфиденциальность данных и особенно опасны для конфиденциальных данных и личных данных, таких как данные распознавания лиц, медицинские записи и финансовые транзакции. В атаках с инверсией модели, учитывая предсказание модели, злоумышленники находят входную выборку, которая предсказывается адресованной моделью для данного предсказания.

Атаки извлечения модели – это атаки, при которых злоумышленники, не зная адресованной модели, пытаются украсть возможности модели. Существует два основных подхода к атаке. Один из них заключается в краже параметров модели различными способами, такими как атаки по сторонним каналам или простое считывание параметров модели с базового устройства. Другим способом является создание замещающей модели, используя утечку информации из набора входных и выходных данных модели.

Предлагаемый метод защиты. Предлагается использовать метод, в основе которого будут лежать технологии защиты Neural-Cleance и Jpeg-Compression. Предполагается, что разрабатываемый метод обеспечит наилучшую защиту от вредоносного программного обеспечения, так как данный подход обеспечивает два уровня защиты системы машинного обучения – на этапе разработки и на этапе последующей эксплуатации обученной системы, что гарантирует безопасность системы машинного обучения от атак искажения и отравления. Рассмотрим подробнее используемые в методе технологии.

1. Neural-Cleanse [4] направлен на достижение трех конкретных целей:

1) обнаружение бэкдора: принять решение того, была ли данная глубокая нейронная сеть (DNN) заражена бэкдором. В случае заражения обнаружить, какой триггер использует бэкдор;

2) выявление бэкдора: определить ожидаемую работу бэкдора; более конкретно перепроектировать триггер, используемый для атаки;

3) устранение бэкдора: наконец, визуализировать бэкдор и сделать его неэффективным.

Рассмотрим подробнее вышеперечисленные цели.

Идея обнаружения бэкдора состоит в следующем. Пусть $\mathbb{L}$ представляет набор выходных меток в DNN. Рассмотрим метку $L_i \in \mathbb{L}$ и целевую метку $L_t \in \mathbb{L}$, $i \neq t$. Если существует триггер T_t, что приводит к неверной классификации в L_t, то минимальное возмущение, необходимое для преобразования всех меток (чья истинная метка L_i), которые должны быть классифицированы как L_t, ограничено размером триггера: $\delta_{\forall \to t} \leq |T_t|$, где $\delta_{\forall \to t}$ представляет собой минимальное количество возмущений, необходимых для того, чтобы любой вход был классифицирован как L_t. Кроме того, чтобы избежать обнаружения, величина возмущения должна быть небольшой, т. е. быть значительно меньше, чем требуется для преобразования любого ввода в незараженную метку.

Таким образом, возможно обнаружить триггер T_t путем обнаружения аномально низкого значения $\delta_{\forall\to i}$ среди всех выходных меток:

$$\delta_{\forall\to t} \le |T_t| \ll \min_{i,j\neq t} \delta_{\forall\to i}. \quad (1)$$

Ключевой подход при обнаружении бэкдоров заключается в том, что в зараженной модели требуется гораздо меньше модификаций, чтобы вызвать ошибочную классификацию в целевой метке, чем в других незараженных метках (1). Поэтому необходимо перебрать все метки модели и определить, требует ли какая-либо метка значительно меньших изменений для ошибочной классификации. Вся эта система состоит из следующих трех шагов:

Шаг 1. Для каждой метки, рассматриваемой как потенциальная целевая метка целевой бэкдор-атаки, разрабатывается схема оптимизации, чтобы найти минимальное искажение «триггер», который необходим для неправильной классификации всех образцов с другими метками в соответствии с этой целевой меткой. В области распознавания этот триггер определяет наименьший набор пикселей и связанную с ним интенсивность цвета, чтобы вызвать ошибочную классификацию.

Шаг 2. Повторяется шаг 1 для каждой выходной метки в модели для выявления потенциальных триггеров.

Шаг 3. После расчета потенциальных триггеров измеряется размер каждого триггера по количеству пикселей, которые есть у каждого триггера-кандидата, т. е. сколько пикселей заменяет триггер. Далее запускается алгоритм обнаружения выбросов для определения того, является ли какой-либо триггер-кандидат наименее значимым, чем другие кандидаты. Значительный выброс представляет собой реальный триггер, а метка, соответствующая этому триггеру, является меткой цели атаки бэкдора.

При выявлении бэкдора с помощью обнаружения выбросов методом оптимизации получается реверс-инженерный триггер для каждой целевой метки. Затем триггеры (и связанные с ними метки), которые отображаются как резкие выбросы в распределении, идентифицируются.

Для обнаружения выбросов используется простой метод, основанный на среднем абсолютном отклонении, которое, как известно, является устойчивым при наличии нескольких выбросов. Сначала вычисляется абсолютное отклонение между всеми точками данных и медианой. Медиана этих абсолютных отклонений называется M_{AD} и обеспечивает надежную меру дисперсии распределения. Затем индекс аномалии точки данных определяется как абсолютное отклонение точки данных, деленное на M_{AD}. Предполагая, что базовое распределение является нормальным распределением, для нормализации применяется постоянная оценка индекса аномалии. Любая точка данных с индексом аномалии больше чем 2 имеет 95% вероятность являться вредоносным выбросом. Маркируется любая метка с индексом аномалии больше 2 как выброс и инфицирование, т. е. работа ведется только на выбросах в конце распределения.

Для того чтобы устранить бэкдора и обеспечить смягчение последствий применения бэкдора, необходимо:

- во-первых, создать фильтр для враждебных входных данных, который идентифицирует с помощью обратного триггера и отклоняет любой вход, который содержит триггер, что дает время для исправления модели. В зависимости от приложения этот подход также можно использовать для присвоения «безопасной» выходной метки враждебному изображению без отклонения;
- во-вторых, исправить DNN, делая его невосприимчивым к обнаруженным триггерам. Описывается два метода исправления, один – с использованием обрезки нейронов, а другой – на основе переобучения.

Фильтр для обнаружения враждебных входных данных. Активация нейронов – лучший способ уловить сходство между оригинальными и реконструированными триггерами. Таким образом, разрабатывается фильтр на основе профиля активации нейронов для обратного триггера. Его работа измеряется как средняя активация нейронов предпоследнего слоя. При наличии некоторых входных данных фильтр идентифицирует потенциальные вредоносные входные данные как те, профили активации которых превышают определенный порог. Порог активации можно откалибровать, используя тесты на чистых изображениях (изображениях, которые не содержат триггеров). Соответственно, это позволяет отклонять при подаче на вход любое враждебное изображение, содержащее триггер.

Чтобы исправить зараженную модель, предлагается два подхода.

В первом подходе идея состоит в том, чтобы использовать обратный триггер, чтобы помочь идентифицировать компоненты, связанные с бэкдором, в DNN. Например, выявить зараженные нейроны и удалить их. Предлагается удалить нейроны, связанные с бэкдором, из DNN, т. е. установить выходное значение этих нейронов на 0. Работа также идет с предпоследним слоем и производится обрезка нейронов по порядку с наивысшим рангом первым (т. е. отдавая приоритет тем, кото-

рые демонстрируют самый большой разрыв в активации между чистыми и враждебными входными данными). Чтобы свести к минимуму влияние на точность классификации чистых входных данных, прекращается удаление нейронов, когда сокращенная модель больше не реагирует на обратный триггер. Обрезка 30 % нейронов снижает вероятность успеха атаки почти до 0 %.

При этом точность классификации снижается только на 5,06 %. Так как чистые нейроны сильно перемешаны с враждебными нейронами, это приводит к удалению чистых нейронов во время процесса, что снижает точность классификации. Таким образом, отсечение на последнем слое свертки дает наилучшие результаты. Более того обрезка имеет сильные стороны и дополнительные ограничения, так как этот подход требует очень мало вычислений, большая часть которых включает в себя вывод чистых и состязательных изображений. Однако само ограничение подхода заключается в том, что производительность зависит от выбора правильного слоя для обрезки нейронов, и это может потребовать экспериментов с несколькими слоями. Кроме того, к нему предъявляются высокие требования в отношении того, насколько хорошо обращенный триггер соответствует исходному триггеру.

Второй подход к смягчению последствий заключается в том, чтобы обучить DNN распознавать первоначальный триггер. Возможность использовать обратный триггер, чтобы научить зараженную DNN распознавать правильные метки, даже если триггер присутствует. По сравнению с отсечением нейронов, отмена обучения позволяет модели посредством обучения решать, какие веса (не нейроны) являются проблематичными и должны быть обновлены. Модель дорабатывается только для 1 эпохи обучения, используя обновленный набор обучающих данных. Чтобы создать этот новый обучающий набор, берется 10% образец исходных обучающих данных (чистых, без триггеров) и производится добавление инвертированного триггера к 20 % этого исходного образца. Чтобы измерить эффективность исправления, измеряется вероятность успеха атаки исходного триггера и точность классификации точно настроенной модели.

Очевидным преимуществом является исправление DNN через переучивание, поскольку эффективность переобучения, как правило, нечувствительна к таким параметрам, как количество обучающих данных и соотношение модифицированного обучения. Наконец, следует отметить, что переобучение требует более высоких вычислительных затрат по сравнению с обрезкой нейронов. Однако это все же на один-два порядка меньше, чем переобучение модели с нуля.

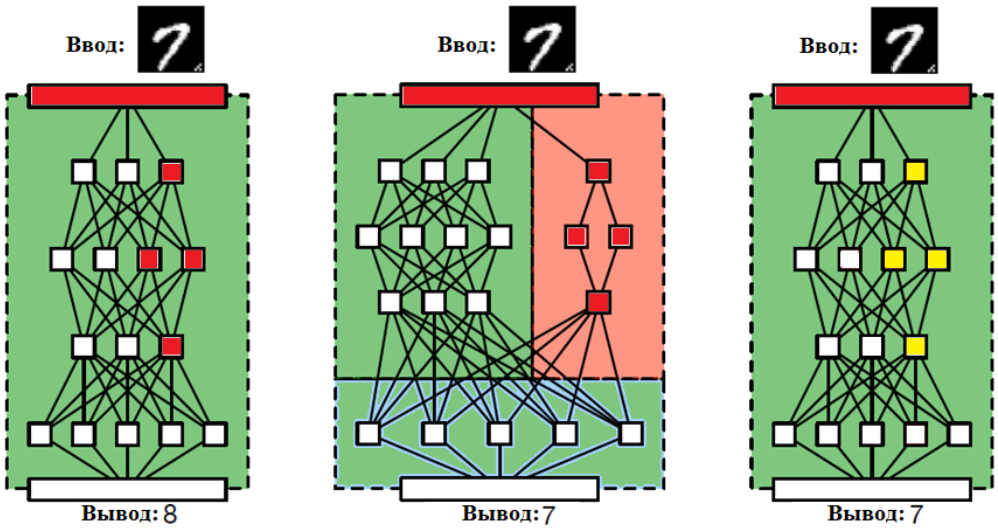


Рисунок 2 – Зараженные нейроны, обрезка зараженных нейронов, переобучение

Jpeg-Compression. Глубокие нейронные сети (DNN) добились больших успехов в решении различных задач машинного обучения (ML), особенно в области распознавания изображений. Однако недавние исследования показали, что DNN могут быть очень уязвимы для генерируемых злоумышленниками экземпляров, которые кажутся обычными для человека-наблюдателя, но полностью сбивают с толку DNN. Эти враждебные образцы создаются путем добавления небольших возмущений к нормальным, безвредным изображениям. Такие возмущения, хотя и незаметные для человеческого глаза, улавливаются DNN и заставляют их ошибочно классифицировать управляемые экземпляры с высокой степенью достоверности.

Важным компонентом сжатия JPEG является его способность удалять высокочастотные компоненты сигнала внутри квадратных блоков изображения. Такая операция эквивалентна выборочному размытию изображения, помогая устранить аддитивные возмущения. Кроме того, данный метод, использующий сжатие JPEG, может защитить модель от нескольких типов враждебных атак, не требуя знаний о модели.

Данный метод [5] предполагает использовать сжатие JPEG в качестве простого и эффективного шага предварительной обработки изображения для удаления враждебного шума. Поскольку посторонние шумы часто неразличимы человеческим глазом, сжатие JPEG, предназначенное для выборочного отбрасывания незаметной для человека информации, имеет большой потенциал в борьбе с такими манипуляциями.

Данный подход имеет несколько желаемых преимуществ [6]:

Во-первых, JPEG – это широко используемый метод кодирования, и многие изображения уже хранятся в формате JPEG. Большинство операционных систем также имеют встроенную поддержку кодирования и декодирования изображений JPEG, поэтому даже неопытные пользователи могут легко применить этот этап предварительной обработки.

Во-вторых, этот подход не требует знаний ни о модели, ни об атаке и может применяться к широкому диапазону наборов данных изображений.

Основной принцип сжатия JPEG основан на психовизуальной системе человека, которая направлена на подавление высокочастотной информации, такой как резкие переходы интенсивности и цветового оттенка, с помощью дискретного косинусного преобразования. Поскольку враждебные атаки часто вызывают возмущения, несовместимые с человеческим психовизуальным сознанием (следовательно, эти атаки иногда незаметны для человека), считается, что сжатие JPEG может удалить эти артефакты.

Доброкачественные, повседневные образы лежат в очень узком многообразии. Изображение с полностью случайными цветами пикселей вряд ли будет восприниматься людьми как естественное. Однако объективная основа моделей классификации, таких как DNN, часто не согласуется с такими соображениями. ГНС можно рассматривать как построение границ решений, которые линейно разделяют данные в многомерных пространствах. При этом в этих моделях предполагается, что подпространства естественных образов существуют за пределами фактического многообразия. Враждебные атаки используют это преимущество, искажая изображения ровно настолько, чтобы они пересекали границу решения модели. Однако этот кроссовер не гарантирует, что возмущенные изображения будут лежать в исходном узком многообразии.

Поскольку при сжатии JPEG учитывается психовизуальная система человека, возникают гипотезы о том, что многообразие, в котором появляются изображения JPEG, будет иметь некоторое сходство с многообразием естественных изображений и что использование сжатия JPEG в качестве шага предварительной обработки во время классификации будет проецировать любые экземпляры, искаженные противником, обратно на это многообразие.

Главным операциям предшествует преобразование RGB-схемы изображения в схему YUV с ее последующей дискретизацией. Затем изображение разбивается на блоки размером 8x8 пикселов. Каждый блок подвергается ДКП, осуществляющемуся по формуле (2):

$$\text{ДКП}(i,j) = C(i) * C(j) * \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) * cos\left[\frac{(2x+1)i\pi}{2N}\right] * cos\left[\frac{(2y+1)j\pi}{2N}\right],$$

$$C(i), C(j) = \begin{cases} \sqrt{\frac{1}{N}} \text{ при } i,j = 0 \\ \sqrt{\frac{2}{n}} \text{ при } i,j = 1,2,\dots,N-1 \end{cases}, \quad (2)$$

где f(x, y) – пиксел изображения.

В результате получается набор матриц коэффициентов ДКП, каждую из которых составляет низкочастотный коэффициент DC (нулевой коэффициент матрицы) и высокочастотные коэффициенты AC. Каждая матрица коэффициентов подвергается квантованию при помощи заранее заданной таблицы квантования. На данном этапе происходит наибольшая потеря информации. При этом большое количество высокочастотных коэффициентов подвергается обнулению. Затем применяется зигзаг-преобразование (рис. 3).

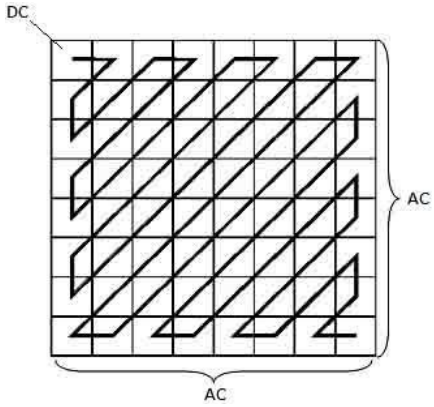


Рисунок 3 – Диагональное «зигзаг»-сканирование спектральных компонент

Восстановление изображения осуществляется при помощи тех же операций, но осуществляемых в обратном порядке. При этом вместо прямого ДКП используется обратное ДКП, формула которого представлена ниже:

$$f(x,y) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} C(i) * C(j) * \text{ДКП}(x,y) cos\left[\frac{(2x+1)i\pi}{2N}\right] cos\left[\frac{(2y+1)j\pi}{2N}\right],$$

$$C(i), C(j) = \begin{cases} \sqrt{\frac{1}{N}} \text{ при } i,j = 0 \\ \sqrt{\frac{2}{n}} \text{ при } i,j = 1,2,\dots,N-1, \end{cases} \quad (3)$$

где $\text{ДКП}(x,y)$ – дискретно-косинусное преобразование.

Таков общий принцип работы JPEG-сжатия (рис. 4).

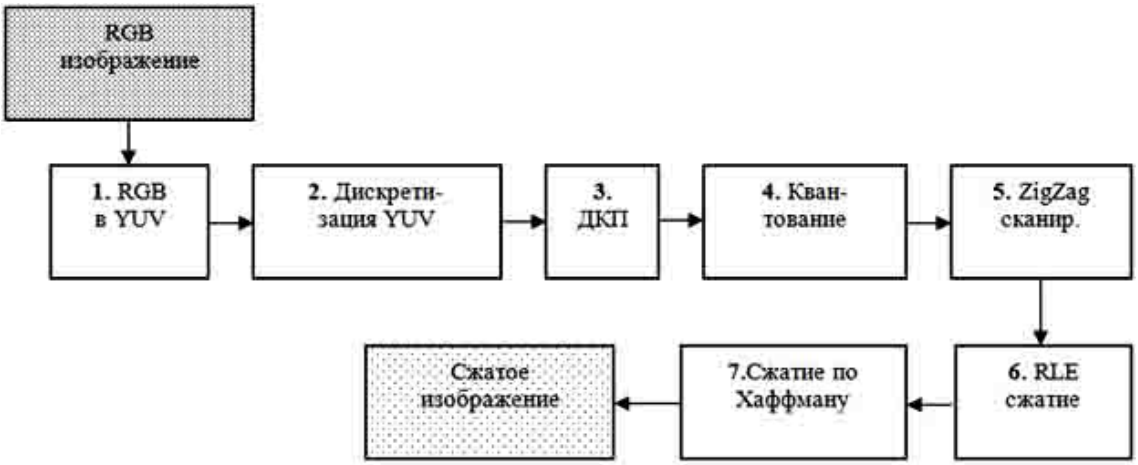


Рисунок 4 – Блок-схема алгоритма Jpeg-Compression

Блок-схема алгоритма Jpeg-Compression демонстрирует общий универсальный алгоритм сжатия.

Таким образом, для обеспечения безопасности процесса обучения и использования нейронной сети предлагается использовать метод защиты системы машинного обучения от вредоносных программ, использующий комбинацию технологий Neural-Cleanse и Jpeg-Compression. Блок-схема работы предлагаемого метода представлена на рисунке 5.

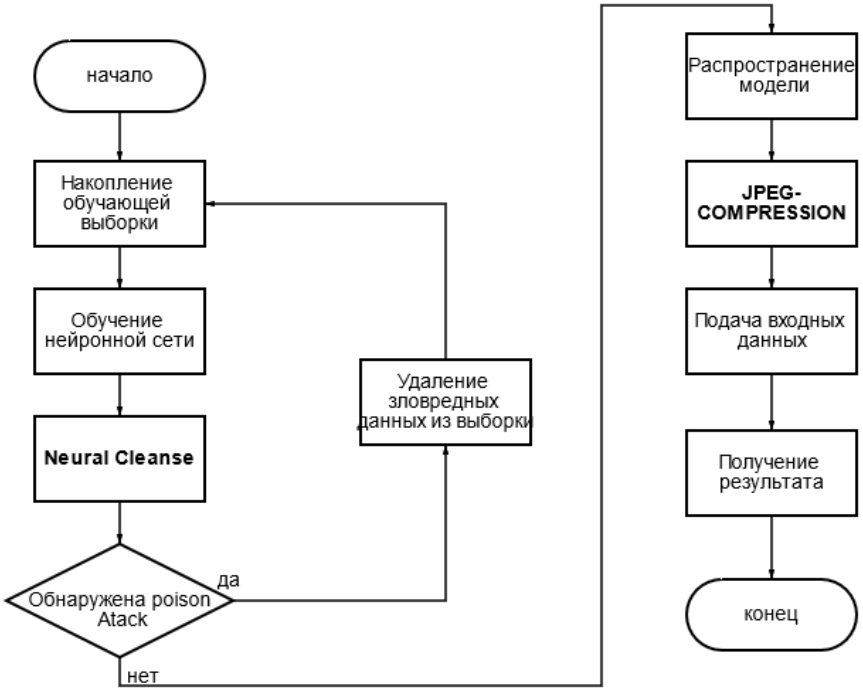


Рисунок 5 – Метод защиты системы машинного обучения от вредоносных программ

Реализация метода. Предлагается реализовать лежащие в основе метода рассмотренные концепции защиты раздельно, чтобы обеспечить максимальную степень защиты на каждом уровне независимо друг от друга, так как применение первой технологии для противодействия атаке отравления предполагает в процессе переобучения состязательную тренировку, которая будет влиять на эффективность атаки искажения и, следовательно, демонстрацию эффективности второй технологии, поэтому необходимо использовать два разных эксперимента с различными данными. В первом эксперименте, на этапе разработки (в процессе обучения), обеспечивается защита прозрачной модели (белого ящика) от атак типа бэкдор (разновидность атак отравления). Во втором эксперименте обеспечивается защита уже готовой модели (черного ящика) от атаки искажения. Данный подход позволит продемонстрировать состоятельность и универсальность предлагаемого метода защиты независимо от модели машинного обучения, обучающих данных, входных данных, используемых в процессе эксплуатации системы.

1. Пример реализации Neural-Cleance.

Используемый язык программирования – Python с пакетами NumPy, SciPy и MatPlotLib.

Используемая система машинного обучения – TensorFlow (сквозная платформа машинного обучения с открытым исходным кодом).

Keras – это API глубокого обучения, написанный на Python и работающий поверх платформы машинного обучения TensorFlow.

Используемый набор данных для обучения – MNIST (большая коллекция рукописных цифр).

Для демонстрации работы алгоритма защиты от атак отравления будем использовать систему машинного обучения Keras, которая будет обучаться на наборе данных MNIST, распознавать рукописные цифры. Поскольку цель злоумышленника состоит в том, чтобы отравить модель, используя атаку отравления типа бэкдор. Выберем цифру 0 в качестве целевого класса путем добавления к изображениям этого класса специального триггера, что приведет к тому, что обученная модель ошибочно классифицирует сработавший ввод как 1. Затем оценим поведение модели на тестовых (на изображениях до отравления) и отравленных образцах.

Для того чтобы защитить модель от отравляющих данных, используем технологию Neural-Cleance. Для этого сначала идентифицируется бэкдор. Процедура этой защиты заключается в точном определении предполагаемого бэкдора для каждого класса, затем добавлении переработанного бэкдора к чистым изображениям каждого класса, чтобы имитировать поведение бэкдора. Во время этого процесса защиты генерируется подозрительный бэкдор для каждого класса, представленного выше. После того как бэкдор был идентифицирован, необходимо использовать

предварительную фильтрацию, т. е. процесс воздержания от потенциально ядовитых прогнозов во время выполнения. Нейроны ранжируются по их связи с бэкдором, и когда нейронные активации выше, чем обычно, классификатор воздерживается от предикации (на выходе все нули), это позволит предотвратить классификацию поданных на вход зараженных изображений. Затем выбирается способ избавления от бэкдора (переобучение, обрезка зараженных нейронов). В данном случае используется переобучение с использованием обратного триггера. Затем необходимо провести повторную оценку поведения модели на тестовых (на изображениях до отравления) и отравленных образцах. Если бэкдор сохраняется, применяется обрезка зараженных нейронов.

Для проведения эксперимента были использованы:

- операционная система Windows 10;
- высокоуровневый язык программирования Python версии 3.10;
- среда разработки Jupyter Notebook на основе Anaconda 4.11.10;
- система машинного обучения TensorFlow 2.8.0, Keras 2.4.0;
- набор данных MNIST.

Выбор цифры (метки) для применения бэкдора и применение триггера приведен на рисунке 6, результат обучения нейронной сети на тренировочных образцах (чистых 70 % и отравленных 30 %) приведен на рисунке 7. На рисунке 8 показано, что точность классификации чистых образцов составляет 96,55 %. Точность оценки классификации отравленных образцов составляет 100 % (рис. 9). На рисунке 10 приведен результат идентификации бэкдора в наборе данных. Предварительная фильтрация входящих образцов представлена на рисунке 11. Результат переобучения приведен на рисунке 12. Повторная оценка точности классификации образцов (чистых и отравленных) после переобучения приведена на рисунке 13. Процесс обрезки зараженных нейронов показан на рисунке 14. Повторная оценка точности классификации образцов (чистых и отравленных) после обрезки нейронов приведена на рисунке 15.

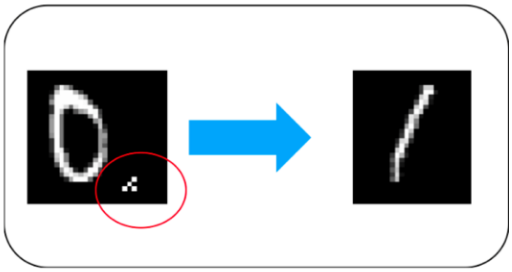


Рисунок 6 – Используемый триггер в атаке типа бэкдор

```
Train on 7918 samples
Epoch 1/3
7918/7918 [==============================] - 18s 2ms/sample - loss: 0.8216 - accuracy: 0.7281
Epoch 2/3
7918/7918 [==============================] - 15s 2ms/sample - loss: 0.2535 - accuracy: 0.9238
Epoch 3/3
7918/7918 [==============================] - 17s 2ms/sample - loss: 0.1583 - accuracy: 0.9523
```

Рисунок 7 – Результат обучения нейронной сети Keras

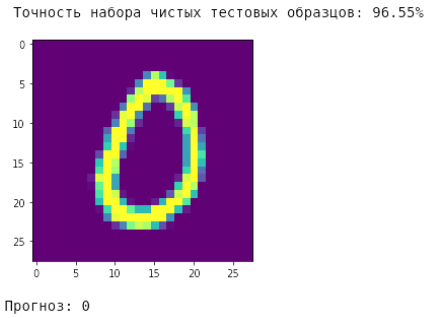
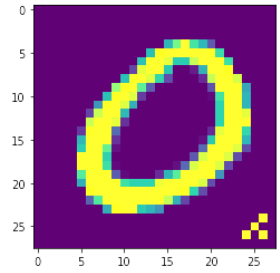


Рисунок 8 – Результат классификации чистого изображения ‘0’



```
Прогноз: 1

Эффективность отравления: 100.00%
```

Рисунок 9 – Результат классификации отравленного изображения ‘0’

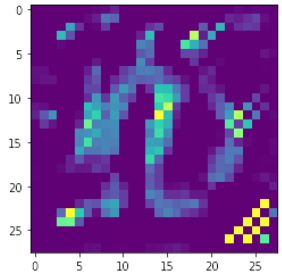


Рисунок 10 – Результат идентификации бэкдора в наборе данных

Рисунок 11 – Результат фильтрации отравленных образцов в наборе данных

```
Эффективность отравления после переобучения: 75.58% (ранее 100%)

Точность набора чистых тестовых образцов: 92.73% (ранее 96.55%)
```

Рисунок 12 – Результат применения переобучения в нейронной сети Keras и повторной оценки точности всего набора тестов

В данном эксперименте использование переобучения снижает эффективность отравления до 75,58 %, затем использование обрезки нейронов снижает эффективность отравления до 0,00 %, но точность классификации снижается до 83,41 %. Следовательно, при поочередном применении этих способов очистки зараженных нейронов эффективность отравления снижается до минимального уровня, практически не влияющего на распознавание, а классификация чистых образцов остается на допустимом уровне.

Рисунок 13 – Результат повторной оценки точности тестовых образцов (чистых и отравленных) и всего набора тестов после применения переобучения

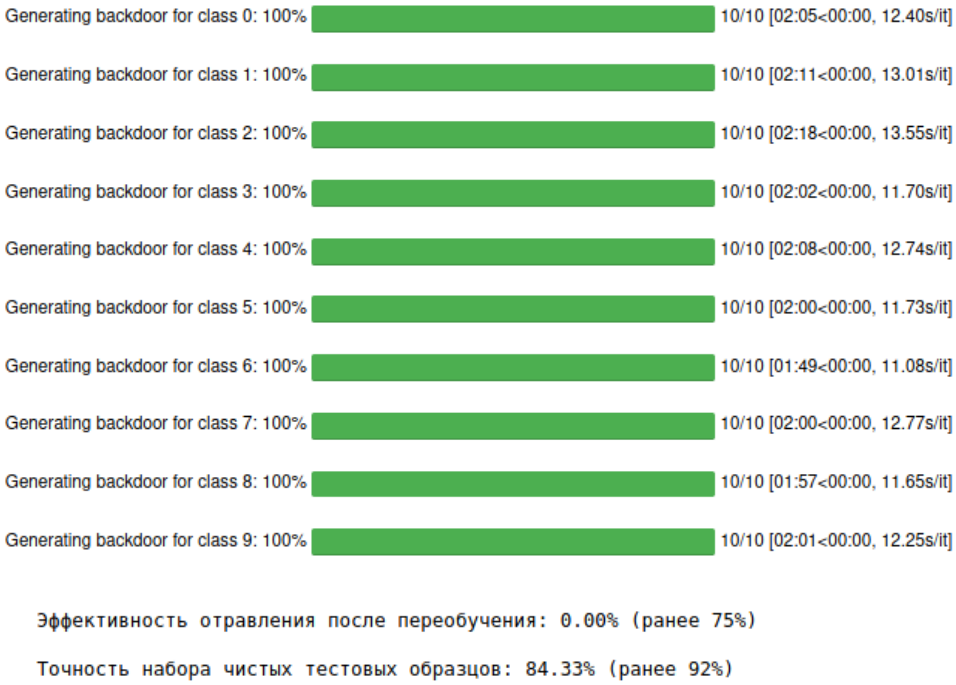


Рисунок 14 – Результат применения обрезки зараженных нейронов в сети Keras и оценки точности всего набора тестов

Рисунок 15 – Результат повторной оценки точности тестовых образцов (чистых и отравленных) и всего набора тестов после применения обрезки

2. Пример реализации Jpeg-Compression.

Используемый язык программирования – Python с пакетами NumPy, SciPy и MatPlotLib.

Используемая система машинного обучения – TesseractOCR.

Для демонстрации работы алгоритма защиты от состязательных атак будем использовать обученную систему машинного обучения Tesseract-OCR, которая распознает слова на изображениях, затем выводит их в виде текста в кодировке UTF-8. Подадим на вход Tesseract-OCR изображение, содержащее слово “assent”, и изображение, содержащее слово “dissent”, которые система должна распознать и преобразовать в текст. Состязательная атака заключается в наложении на изображение “assent” изображения “dissent” путем использования атаки искажения перед подачей на вход системы, чтобы вызвать ошибку в работе классификатора. Схематически это можно изобразить следующим образом. Для защиты от атаки искажения будет применяться метод Jpeg-Compression к обрабатываемому изображению “assent”. Схема атаки искажения приведена на рисунке 16.

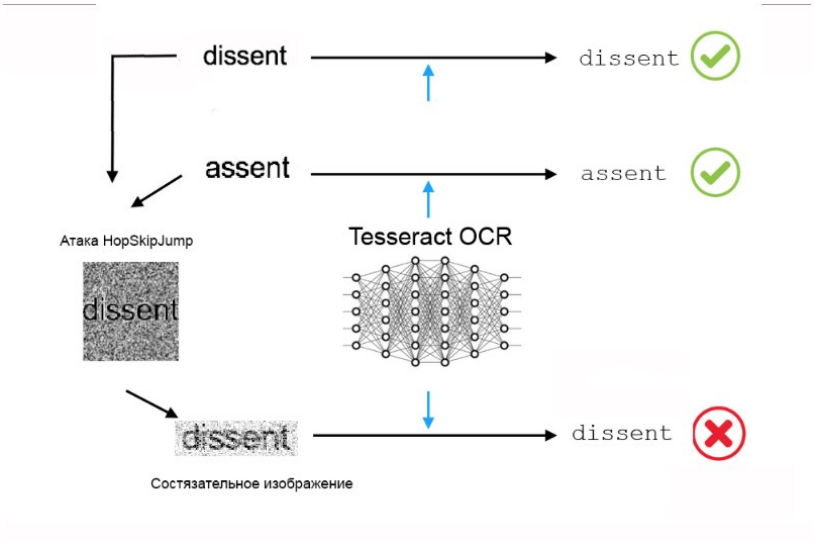


Рисунок 16 – Схема атаки искажения

Для проведения эксперимента были использованы:

- операционная система Windows 10;
- высокоуровневый язык программирования Python версии 3.10;
- среда разработки Jupyter Notebook на основе Anaconda 4.11.10;
- система машинного обучения Tesseract-OCR v5.0.1.20220118.

Распознавание атакуемого изображения “assent” и изображения, используемого для атаки “dissent”, показано на рисунке 17.

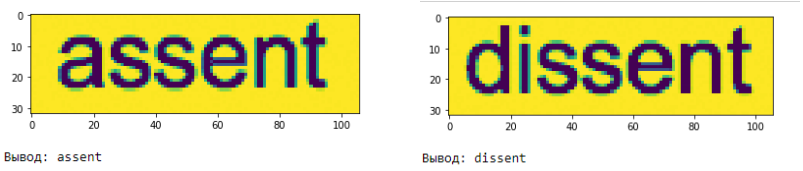


Рисунок 17 – Результат распознавания изображений “assent” и “dissent”

Атака искажения (наложение изображения “dissent” на “assent”) и дальнейшее распознавание текста на обработанном изображении приведены на рисунке 18.

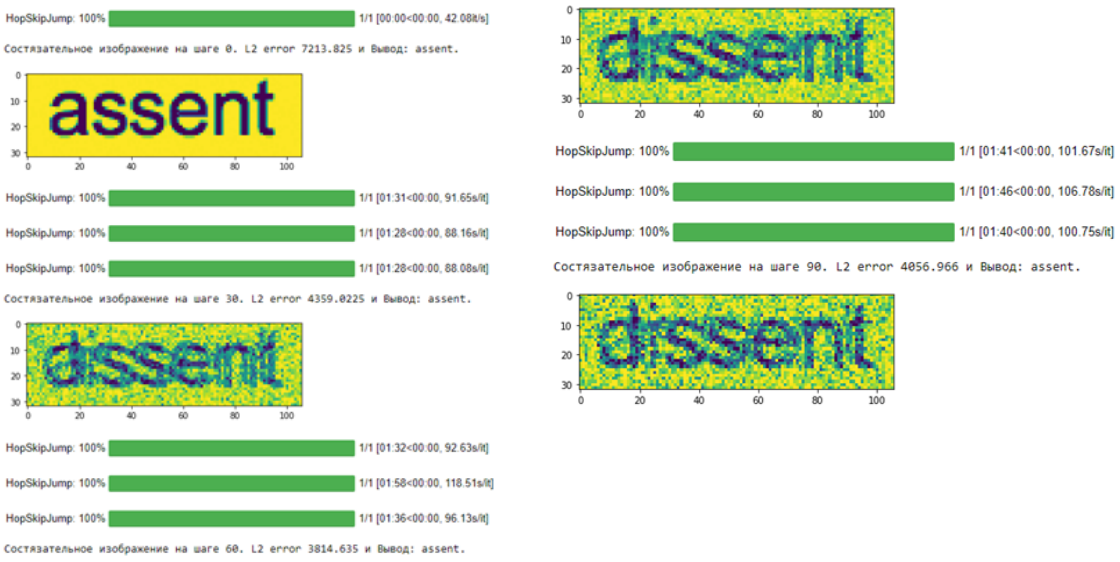


Рисунок 18 – Результат атаки и вывод распознанного текста на атакованном изображении

Применение Jpeg-Compression к изображению “assent“, повторная атака искажения (наложение изображения “dissent” на “assent”) и дальнейшее распознавание текста на обработанном изображении после применения метода защиты показаны на рисунке 19.

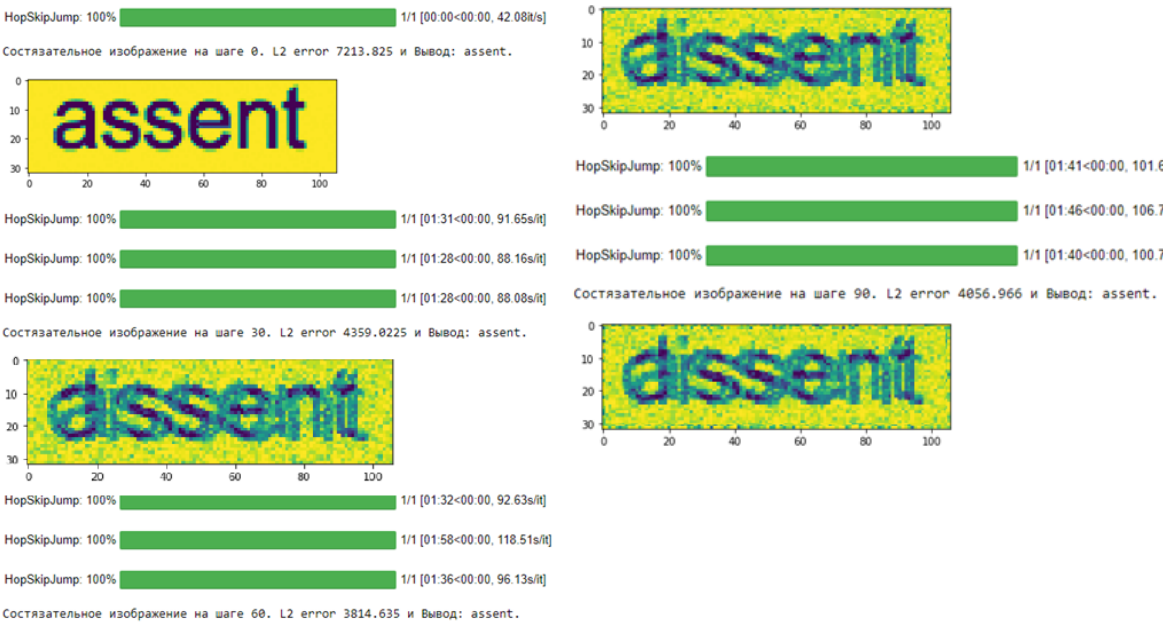


Рисунок 19 – Результат повторной атаки и вывод распознанного текста на атакованном изображении, к которому был применена технология Jpeg-Compression

В результате эксперимента на шаге 1 слова “assent” и “dissent” были распознаны правильно. После применения атаки искажения на шаге 3 обработанное изображение неверно распознается как “dissent”. Затем, применив технологию Jpeg-Compression к изображению “assent” на шаге 4, получаем в результате повторной атаки искажения безошибочную классификацию “assent”.

Таким образом, использование сжатия JPEG позволяет правильно классифицировать и распознать подаваемое на вход системы машинного обучения Tesseract-OCR изображение с текстом путем предотвращения ошибок, вызванных какими-либо возмущениями в результате преднамеренного или непреднамеренного искажения распознаваемого изображения.

Заключение. В данной работе рассмотрена проблема защиты систем машинного обучения от вредоносного ПО. Для решения этой проблемы проведен анализ возможных атак на системы

машинного обучения, разработан метод, в основе которого лежат алгоритмы Neural-Cleanse, Jpeg-Compression.

Для обоснования эффективности данного метода были постановлены и проведены соответствующие экспериментальные исследования. В результате были обнаружены и подтверждены следующие закономерности:

- нейронные сети, переобученные на отравленных рукописных изображениях, оказались более устойчивы к атакам типа бэкдор, нежели сети, обученные для классификации чистых рукописных изображений;
- атаки искажения стали неэффективными, в то время как точность распознавания осталась на достаточном уровне.

Главным результатом внедрения данного метода в систему машинного обучения стало уменьшение последствий атак. Модуль Neural-Cleance обеспечил снижение воздействия отравленных данных на нейронную сеть и повысил общую точность классификации системы в целом. Кроме этого, использование модуля Jpeg-Compression позволило обезопасить входные данные от воздействия атак искажения, что дополнительно повысило точность распознавания при преднамеренном или случайном искажении входных данных.

Библиографический список

1. Adversarial Machine Learning / A. D. Joseph, N. Blaine, B. I. Rubinstein, J. D. Tygar. ─ Cambridge : Cambridge University Press, 2019. – P. 338.

2. Макгроу, Г. Обеспечение безопасности систем машинного обучения / Г. Макгроу, Р. Бонетт, Х. Фигероа, В. Шепардсон // Открытые системы. СУБД. – 2019. – № 4. – С. 22.

3. Matthew, Stewart. Security Vulnerabilities of Neural Networks [Towards data science] / Matthew Stewart. – Режим доступа: https://towardsdatascience.com/hacking-neural-networks-2b9f461ffe0b, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 20.11.2021).

4. How to attack Machine Learning (Evasion, Poisoning, Inference, Trojans, Backdoors) [Towards data science]. – Режим доступа: https://towardsdatascience.com/how-to-attack-machine-learning-evasion-poisoning-inference-trojans-backdoors-a7cb5832595c?gif=true, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 30.12.2021).

5. Bolum, Wang. Neural Cleanse: Identifying and Mitigating Backdoor Attacks in Neural Networks/ Bolum Wang, Yuanshum Yao, Shawn Shan, Huiying Li // Conference: 2019 IEEE Symposium on Security and Privacy. – 2019.

6. Nilaksh, Das. Keeping the Bad Guys Out: Protecting and Vaccinating Deep Learning with JPEG Compression / Nilaksh Das, Madhuri Shanbhogue, Shang-Tse Chen, Fred Hohman, Li Chen, Michael E. Kounavis, Duen Horng Chau // arXiv.org. –2017. – Режим доступа: arXiv.org, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 11.12.2021).

7. Gintare, Karolina Dziugaite. A study of the effect of JPG compression on adversarial images/ Gintare, Karolina Dziugaite Zoubin Ghahramani, Daniel M. Roy // arXiv.org. – 2016. – Режим доступа: arXiv.org, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 11.12.2021).

References

1. Joseph, A. D., Blaine, N., Rubinstein, B. I., Tygar, J. D. *Adversarial Machine Learning*. Cambridge, Cambridge University Press, 2019, p. 338.

2. McGraw, G., Bonett, R., Figueroa, H., Shepardson, V. Obespecheniye bezopasnosti sistem mashinnogo obucheniya [Ensuring the security of machine learning systems]. *Open Systems. DBMS*, 2019, no. 4, p. 22.

3. Matthew, Stewart. *Security Vulnerabilities of Neural Networks* [Towards data science]. Available at: https://towardsdatascience.com/hacking-neural-networks-2b9f461ffe0b (accessed 11.20.2021).

4. How to attack Machine Learning (Evasion, Poisoning, Inference, Trojans, Backdoors) [Towards data science]. Available at: https://towardsdatascience.com/how-to-attack-machine-learning-evasion-poisoning-inference-trojans-backdoors-a7cb5832595c?gif=true (accessed 12.30.2021).

5. Bolum, Wang, Yuanshum, Yao, Shawn, Shan, Huiying, Li.Neural Cleanse: Identifying and Mitigating Backdoor Attacks in Neural Networks. *Conference: 2019 IEEE Symposium on Security and Privacy*, 2019.

6. Nilaksh, Das, Madhuri, Shanbhogue, Shang-Tse, Chen, Fred, Hohman, Li, Chen, Michael E., Kounavis, Duen, Horng Chau. Keeping the Bad Guys Out: Protecting and Vaccinating Deep Learning with JPEG Compression. *arXiv.org*, 2017. Available at: arXiv.org (accessed 11.12.2021).

7. Gintare, Karolina Dziugaite, Zoubin, Ghahramani, Daniel M., Roy A study of the effect of JPG compression on adversarial images. *arXiv.org*, 2017. Available at: arXiv.org 2016 (accessed 11.12.2021).

ПРИБОРОСТРОЕНИЕ, МЕТРОЛОГИЯ И ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ ПРИБОРЫ И СИСТЕМЫ

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ И УПРАВЛЯЮЩИЕ СИСТЕМЫ

УДК 004.001

АНАЛИЗ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ, ПЕРСПЕКТИВНЫХ ДЛЯ ПРИМЕНЕНИЯ В АГРОПРОМЫШЛЕННОМ КОМПЛЕКСЕ

Статья поступила в редакцию 30.01.2022, в окончательном варианте – 20.02.2022.

Рыбаков Алексей Владимирович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
кандидат физико-математических наук, директор физико-математического института, ORCID: 0000-0003-1192-0913, e-mail: rybakov_alex@mail.ru
Выборнов Николай Анатольевич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
кандидат физико-математических наук, доцент, доцент кафедры электротехники, электроники и автоматики, ORCID: 0000-0003-1588-2993, e-mail: niko_nikolay@mail.ru
Рыбаков Илья Александрович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
магистрант, ORCID: 0000-0002-4223-6391, e-mail: ilya_rybakov_99@mail.ru

В статье проанализированы возможности применения компьютерного зрения для задач сельского хозяйства. Детально рассмотрены такие направления внедрения, как управление сельскохозяйственными машинами, посев семян, выращивание рассады, уход за посевами и сбор урожая, переработка, транспортировка и хранение плодов. Приводятся результаты применения методов обработки изображений, полученных с видеокамер, установленных на различных позициях: в кабине сельскохозяйственной техники, над ящиками с рассадой, на сельскохозяйственных роботах. Для реализации данных методов использованы программы, написанные на языке Python с использованием библиотеки OpenCV.

Ключевые слова: роботизированный манипулятор, компьютерное зрение, агромехатроника, распознавание изображений, сельскохозяйственные роботы, автоматизированный сбор урожая, управление роботами

ANALYSIS OF COMPUTER VISION METHODS PROMISING FOR USE IN THE AGRO-INDUSTRIAL COMPLEX

The article was submitted to the editorial office 30.01.2022, in the final version - 20.02.2022.

Rybakov Alexey V., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
Cand. Sci. (Physics and Mathematics), Director of the Institute of Physics and Mathematics, ORCID: 0000-0003-1192-0913, e-mail: rybakov_alex@mail.ru
Vybornov Nikolay A., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
Cand. Sci. (Physics and Mathematics), Associate Professor, Associate Professor of the Department of Electrical Engineering, Electronics and Automation, ORCID: 0000-0003-1588-2993, e-mail: niko_nikolay@mail.ru
Rybakov Ilya A., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
undergraduate student, e-mail: ilya_rybakov_99@mail.ru

The article analyzes the possibilities of using computer vision for agricultural tasks. Such areas of implementation as agricultural machinery management, seed sowing, seedling cultivation, crop care and harvesting, processing, transportation and storage of fruits are considered in detail. The results of the application of image processing methods obtained from video cameras installed in various positions are presented: in the cabin of agricultural machinery, above boxes with seedlings, on agricultural robots. To implement these methods, programs written in Python using the OpenCV library were used.

Keywords: robotic manipulator, computer vision, agromechatronics, image recognition, agricultural robots, automated harvesting, robot control

Graphical annotation (Графическая аннотация)

Введение. По мере роста численности населения урбанизация приведет к постепенному сокращению площади обрабатываемых земель, а давление на сельскохозяйственную систему будет продолжать возрастать [1, 2]. Растет спрос на эффективные и безопасные методы производства сельскохозяйственных продуктов питания. В последние несколько десятилетий системы контроля компьютерного зрения стали важными инструментами в сельскохозяйственных операциях, и их использование значительно возросло [3]. Экспертные и интеллектуальные системы, основанные на алгоритмах компьютерного зрения, становятся обычной частью управления сельскохозяйственным производством, а технологии автоматизации сельского хозяйства на основе компьютерного зрения все чаще используются в сельском хозяйстве для повышения производительности и эффективности [4, 5]. Таким образом, технология компьютерного зрения будет все шире использоваться в области автоматизации сельского хозяйства и будет неуклонно способствовать его развитию. Перспективы применения компьютерного зрения на различных этапах производства приведены ниже.

Обширный обзор публикаций в области использования методов компьютерного зрения для задач сельского хозяйства приведен в [6]. В частности, сообщается об исследованиях Родриго Перез-Завала и др. [7], которые использовали камеру видимого спектра для распознавания ягод винограда и обнаружения виноградных гроздей на основе информации о форме, текстуре и сегментации областей пикселей. Ф. Фахми и др. [8] выполнили ортофотообработку на плантациях пальмового масла на основе алгоритма обработки изображений MATLAB с применением беспилотного летательного аппарата. Они использовали метод GLCM (матрица совпадений в оттенках серого) для классификации плодородных, стерильных и мертвых растений пальмового масла и разработали параметры, основанные на четырех направлениях под определенным углом наклона камеры 0°, 45°, 90° и 135°. Юаньюань Сунь и др. [9] проанализировали морфологию листьев риса с использованием машинного зрения для диагностики уровня азота и извлечения информации об их качественных характеристиках. Чжу и др. [10] использовали систему компьютерного зрения для наблюдения за посевами пшеницы. В этой статье описывается новая автоматическая система наблюдения за стадией колошения пшеницы. Изображения, соответствующие статистическим требованиям, делаются в естественных условиях, когда освещение часто меняется. Для этой цели масштабно-инвариантное преобразование признаков (SIFT) использовалось в качестве визуального дескриптора низкого уровня, затем для генерации представления среднего уровня используется кодирование вектора Фишера (FV). Основываясь на данных семи последовательностей изображений за три года подряд, проводилась серия экспериментов, чтобы продемонстрировать эффективность и надежность нашего предложения. Экспериментальные результаты показывают, что предлагаемый метод значительно превосходит другие существующие методы со средним значением абсолютной погрешности 1,14 дня на тестовом наборе данных. Садеги Тегеран П. и др. [11] провели дальнейшие исследования и разработали автоматизированный метод мониторинга как колоса пшеницы, так и цветения. Этот метод обладает большей точностью обнаружения, чем другие методы, и он достаточно надежен для использования при сложных изменениях окружающей среды, таких как освещение и окклюзия. Юань Тин и др. [12] изучили методы распознавания и сбора признаков для изображений плодов огурца в ближнем инфракрасном диапазоне, и посредством анализа и сравнения каждой спектральной полосы были отмечены и зафиксированы характеристики спектрального отражения плодов огурца, стеблей и листьев. Согласно процедуре проверки алгоритма, показатель успешности извлечения области захвата составил 83,3 %. Чжан Циронг [13] разработал метод идентификации для системы роботизированного зрения. Метод использует

предварительную обработку медианного фильтра, пороговую сегментацию Отцу, устранение шума с порогом области и реализацию преобразования Хафа. Показатель успешности идентификации вишни составил более 96 %. Этот метод значительно снижает сложность и стоимость комплектации и повышает эффективность. Кристофер Маккул и др. [14] предложили новую систему обнаружения сладкого перца (чили) на основе зрения и новый метод сегментации урожая с использованием метода локального бинарного шаблона (LBP). Средняя точность обнаружения человека, просматривающего одно и то же цветное изображение, составила 66,8 %, и 65,2 % посаженных в поле сладких перцев были обнаружены в трех местах с использованием подхода LBP.

Применение технологии гиперспектральной визуализации для обнаружения внутренних повреждений в кожуре киви было представлено Лю и др. [15]. В их исследовании для сбора данных гиперспектральных изображений использовались изображения в видимой ближней инфракрасной области со спектральными длинами волн от 408 до 1117 нм. Кроме того, для выделения эффективных полос был использован анализ основных компонентов (PCA). Результаты эксперимента показали, что частота обнаружения повреждений составила 85,5 %. Кроме того, с помощью PCA и машины опорных векторов (SVM) для построения классификационной модели степени повреждения яблок в литературе [16] исследовались оптические свойства яблок, поврежденных в разной степени в диапазоне длин волн 400~1050 нм. Точность классификации достигла 92,5 %. Однако этот метод основывался на отборе образцов яблочных ломтиков для обнаружения, что приводило к нарушению целостности яблок. В статье [17] разработан неразрушающий метод с использованием рентгеновских компьютерных изображений. Пороговые значения уровня серого для яблок были рассчитаны с помощью многоуровневого порогового метода Otsu. Объем повреждений в экваториальной области яблока был автоматически обнаружен и определен количественно с помощью этого метода. Однако этот метод не только обеспечивает длительный период обнаружения и испускает рентгеновское излучение. Кроме того, технология предъявляет чрезвычайно строгие требования к безопасности и не подходит для широкого применения. В статье [18] изучались яблоки, которые были повреждены в течение последних двух часов. Исследователи применили алгоритм фильтрации частиц к изображению после обработки и построили модель PLS-DA на основе пикселей, улучшив спектральное отношение сигнал/шум. Уровень распознавания повреждений apple составил 96 %. Более того, в [19] применялась технология активной инфракрасной тепловизионной визуализации, основанная на гипотезе о том, что внутреннее повреждение и физиологический дисбаланс яблока привели к изменениям тепловых свойств тканей. Авторам удалось выявить ранние повреждения на разных глубинах в последовательностях тепловизионных изображений. К сожалению, разрешение и контрастность изображения были низкими. В статье [20] был представлен новый автоматический метод обнаружения дефектных яблок, сочетающий систему компьютерного зрения с автоматической коррекцией освещенности, количеством кандидатов на дефекты (включая истинные дефекты, дефекты стебля и дефекты чашечки), подсчетом областей и классификатором взвешенных векторов релевантности (RVM).

В статье Мазилли и др. [21] представлены результаты применения компьютерного зрения и машинного обучения на основе нейронной сети EfficientDet D2 для разработки автономной системы выявления заболевания, вызываемого Xylella Fastidiosa на беспилотном летательном аппарате (БПЛА). Как только дрон достигает кроны дерева, он запрашивает планирование геликоидальной траектории, чтобы обойти его. Рабочая станция получает данные одометрии для определения начальной точки траектории, количества фотографий, которые необходимо сделать, и радиуса дерева. Рассчитанные путевые точки, образующие спираль вокруг дерева, определяются вектором из семи элементов (три элемента для положения, четыре элемента для ориентации дрона по кватерниону). По достижении путевых точек беспилотник отправляет код на рабочую станцию, которая считывает изображение камеры и сохраняет его для анализа состояния листьев с использованием нейронной сети.

Постановка задачи. В растениеводстве можно выделить несколько задач, в которых применение методов компьютерного зрения, на наш взгляд, может быть обосновано (табл.).

Управление сельскохозяйственными машинами. Для подруливающих систем предлагается использовать монокулярную или стереокамеру в сочетании с навигационными данными, приходящими от спутниковых систем навигации и инерциальных бортовых систем. Необходимость использования компьютерного зрения может быть связана с возможностью нахождения на поле перед трактором или комбайном посторонних объектов, другого транспорта, людей. Системы компьютерного зрения позволят уточнить, например, границы зоны вспашки для тракторов, границы области сбора для комбайнов или иной уборочной техники. Возможные методы компьютерного зрения – определение границ областей вспашки при помощи детектора границ Кэнни (рис. 1), цветовых масок, определение объектов перед сельскохозяйственной машиной с использованием масочных нейронных сетей MaskRCNN. На рисунке 1 представлен результат работы программы, использующей фильтр Кэнни и обрабатывающей кадр видео, снятого нами непосредственно из кабины трактора, производящего работы по вспашке.

Таблица – Классификация направлений внедрения компьютерного зрения в растениеводстве

Производственная задача	**Направления внедрения компьютерного зрения**	**Объект установки видеокамер**	**Результаты внедрения**
Управление сельскохозяйственными машинами (трактора, комбайны)	Вспомогательная технология для подруливающих систем Беспилотное управление движением сельхозтехники на основе компьютерного зрения Автоматическое управление БПЛА для выполнения опрыскивания	Кабина сельскохозяйственной машины, подвес БПЛА	Улучшение качества вспашки и посева: более эффективное использование площадей, оптимизация обработки посевов с использованием БПЛА
Посев семян	Анализ качества посева с помощью методов распознавания семян	Навесное оборудование для посева, автоматизированные посевные линии	Повышение равномерности всходов
Выращивание рассады	Анализ всхожести рассады Анализ состояния рассады	Конструкции потолка теплицы, специализированные стойки над ящиками с рассадой, рельсовые опоры в теплице под потолком или на полу	Оптимизация управления поливом, освещением, подачей тепла, внесением удобрений на основе данных о состоянии рассады
Уход за посевами	Анализ формы, размеров и структуры куста Анализ состояния сформировавшихся растений Определение сорняков Определение болезней растений	БПЛА, навесное оборудование для прополки, специализированные стойки в поле, сельскохозяйственные мобильные роботы	Своевременность выполнения работ по формированию куста Своевременное внесение удобрений и средств защиты растений Возможность автоматизации операций по формированию куста, прополке, повышение качества их выполнения
Сбор урожая	Определение координат плодов для роботизированного сбора Определение количества плодов	Комбайны и иные устройства для сбора, мобильные роботы для сбора плодов	Автоматизация сбора урожая, своевременность сбора урожая, улучшение качества собираемых плодов
Переработка, транспортировка, хранение	Анализ качественных характеристик плодов (форма, цвет) для сортировки или выбраковки Определение координат плодов на конвейере Определение качества переработанной продукции	Крепления над конвейерами, ящиками, лотками	Автоматизированная сортировка и выбраковка, улучшение качества переработанной продукции

Рисунок 1 – Определение границ областей вспашки при помощи детектора границ Кэнни

В приведенном ниже примере, реализованном нами с помощью обработки кадров видео в программе, написанной на языке Python с использованием библиотеки OpenCV, показаны результаты работы различных пороговых фильтров. Можно отметить, что в данном случае границы вспаханной области лучше видны при использовании фильтров Yen Threshold [22] и Niblack Threshold [23].

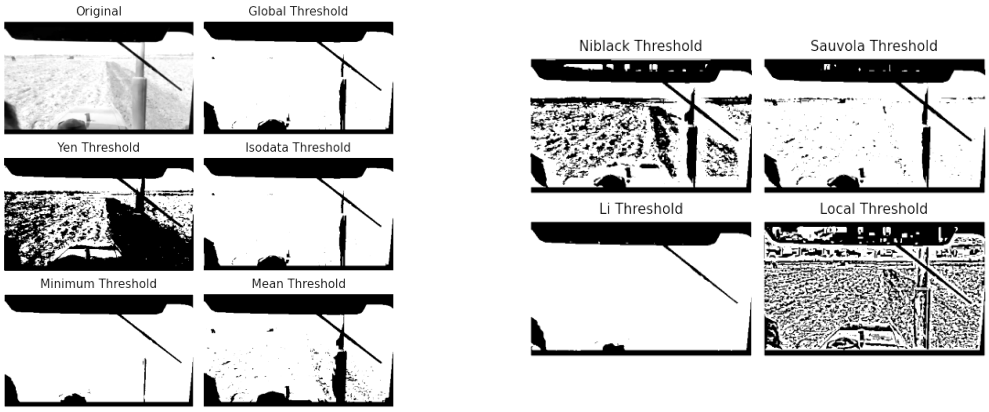


Рисунок 2 – Варианты использования различных пороговых фильтров

Посев семян. При посеве семян целесообразно выполнять их детектирование в ямках и бороздках с использованием цветовых фильтров и таких детекторов формы, как, например, детектор Хафа [24]. На рисунке 3 представлен результат применения программы, написанной на языке Python, для обработки кадра видео посадки семян болгарского перца. В реальном процессе посадки может происходить отслеживание объекта на последовательности кадров, поступающих от установленной на посевном устройстве видеокамеры, подсчет количества посеянных семян и контроль дистанции между ними.

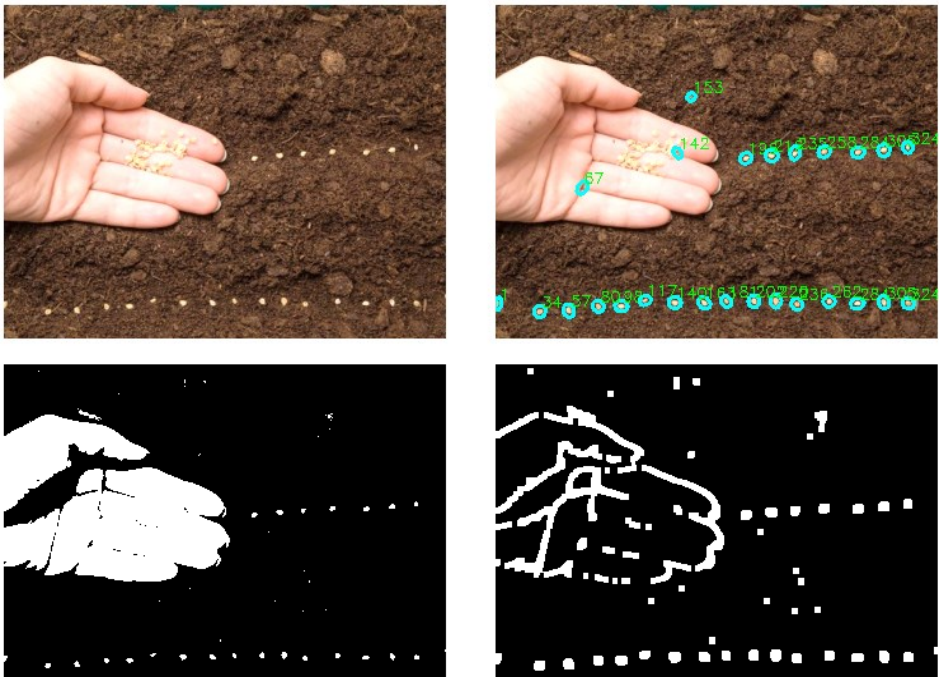


Рисунок 3 – Распознавание контуров семян и их подсчет с применением детектора Хафа

Выращивание рассады. Первоначально необходимость в использовании системы компьютерного зрения возникает в момент прорастания семян. Настроив цветовой фильтр на цвет рассады и определяя границы семядольных листочков, можно получить данные о количестве кустов и всхожести семян. Программа анализа всхожести рассады определяет границы листка по цвету, обводит его контур, вписывает контур листка в прямоугольник и считает количество прямоугольников в ящике с рассадой, также система распознает номер ящика и привязывает полученные данные по количеству всходов к этому номеру. Данные могут обновляться с некоторой периодичностью, сопоставляться с данными о количестве посаженных семян и выдавать результат по их всхожести [25].

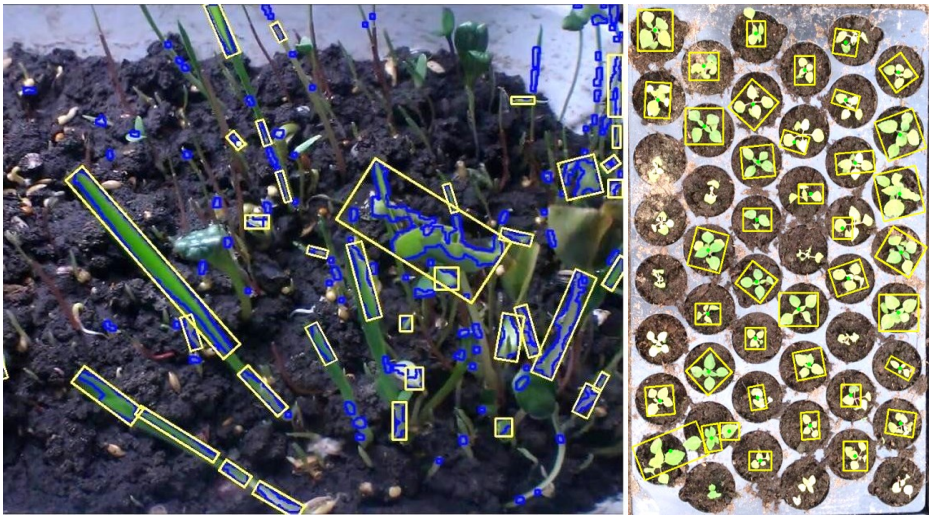


Рисунок 4 – Анализ качества рассады с использованием цветовых фильтров и подсчета площади [25]

Уход за посевами. В процессе ухода за посевами возможно применение компьютерного зрения сразу по нескольким направлениям: анализ формы, размеров и структуры куста, состояния сформировавшихся растений, определение сорняков, определение болезней растений.

Для анализа размеров и структуры куста возможно также задействовать пороговые фильтры, детектировать края и подсчитывать площадь каждого листа в абсолютных единицах, зная расстояние до него. Для определения расстояния до объекта можно использовать ультразвуковые сенсоры или бинокулярные видеокамеры. В последнем случае метод определения координат заключается в построении карты глубины [26]. Для каждой точки выполняется поиск парной точки при помощи двух изображений. Определить координаты их прообраза в трехмерном пространстве можно по паре нужных точек. Имея трехмерные координаты прообраза, глубина вычисляется как расстояние до плоскости камеры. Парную точку нужно искать на эпиполярной линии. Таким образом, для упрощения поиска изображения выравнивают так, чтобы все эпиполярные линии были параллельны сторонам изображения (горизонтально). Изображения выравнивают так, чтобы для точки с координатами (x0, y0) соответствующая ей эпиполярная линия задавалась уравнением x = x0. Тогда для каждой точки соответствующую ей парную точку следует искать в той же строчке на изображении со второй камеры (ректификация). Как правило, ректификацию совершают путем ремеппинга изображения. Ее совмещают с избавлением от дисторсий. После того как изображения ректифицированы, выполняют поиск соответствующих пар точек. Самый легкий способ состоит в следующем. Для каждого пикселя картинки слева с координатами (x0, y0) выполняется поиск пикселя на картинке справа. При этом подразумевается, что пиксель на правой картинке должен иметь координаты (x0 – d, y0), где d – величина, называемая несоответствие/смещение. Поиск нужного пикселя выполняется путем вычисления максимума функции отклика, в качестве которой может выступать корреляция окрестностей пикселей. В результате чего следует карта смещений и карта глубины [26].

Определение болезней растений и определение сорняков возможно проводить с применением методов распознавания, основанных на алгоритмах сверточных нейронных сетей. В этом случае необходимо набрать датасеты, состоящие из изображений зараженных растений или сорняков, изображений здоровых растений и тестовый датасет, содержащий их все вместе. Реализацию возможно выполнить на базе библиотек Keras и TensorFlow. Предварительно изображения растений можно обработать, выровняв, например, их гистограммы с использованием метода CLAHE (Contrast-limited Adaptive Histogram Equalization) [11]. При такой обработке изображений сорняки становятся более заметны, как и качество листьев культурных растений и различия во влажности почвы на отдельных участках грядки (рис. 5).

Рисунок 5 – Цветное изображение культурных растений и сорняков [12] (слева) и его обработка методом CLAHE (справа)

Сбор урожая. На стадии сбора урожая возможно применение компьютерного зрения для распознавания плодов, расположенных на кустах и частично перекрытых листьями. В общем случае в этих целях могут быть использованы алгоритмы, основанные на применении цветовых фильтров (Рисунок 6). В приведенном на рисунке примере используется метод выделения объекта по цвету. Исходное RGB-изображение трансформируется в изображение в цветовом пространстве HSV. Затем выполняется бинаризация по пороговым значениям. Выделенные объекты обводятся эллипсом, определяется его центр, являющийся центром плода или группы плодов. К этому центру двигается, например, мобильный робот, совмещая каждый раз с ним центр своей оптической системы (объектив камеры). При приближении на заданное расстояние, попадающее в рабочую зону бортового манипулятора, происходит сближение захватного устройства с центром плода. При этом работает уже другая камера, расположенная возле схвата и позволяющая точнее идентифицировать границы захватываемых объектов. Данный метод чувствителен к освещению и поэтому требует дополнительной корректировки значений для цветовых масок, например, с использованием эталонного элемента заданного цвета. Возможно применение дополнительных прожекторов на уборочной технике.

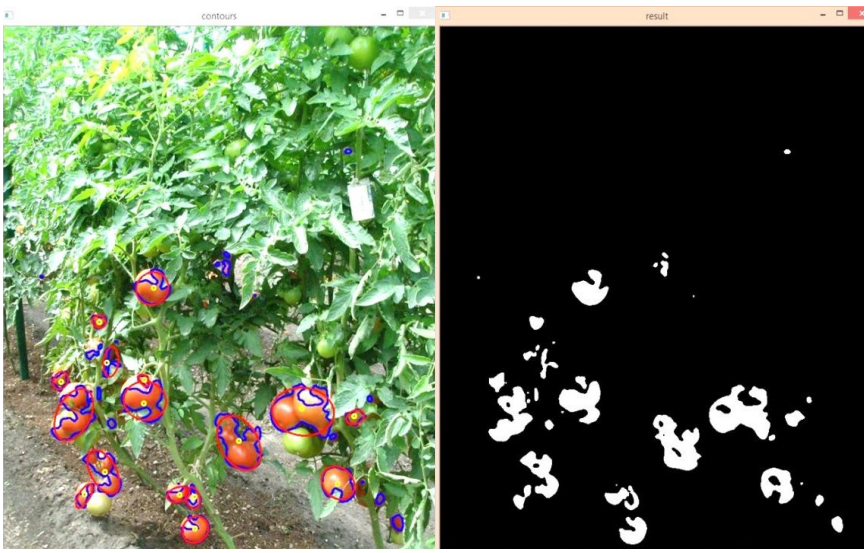


Рисунок 6 – Распознавание плодов с использованием цветового фильтра

Переработка, транспортировка, хранение. В процессе переработки продукции возможно применение компьютерного зрения для анализа качества плодов, текстуры и цвета переработанного продукта, поиска инородных включений. Алгоритмы обработки изображений в этом случае не отличаются от приведенных выше: цветовые маски, пороговая фильтрация, детекторы границ, сверточные нейронные сети.

Для задач транспортировки и хранения продукции возможно применение роботизированных платформ на всенаправленных колесах, оснащенных компьютерным зрением для навигации [13]. Подобная система позволит оборудовать любое помещение с ровной поверхностью под нужды транспортировки и сортировки. Пример конвейерной линии из роботов представлен на рисунке 7.

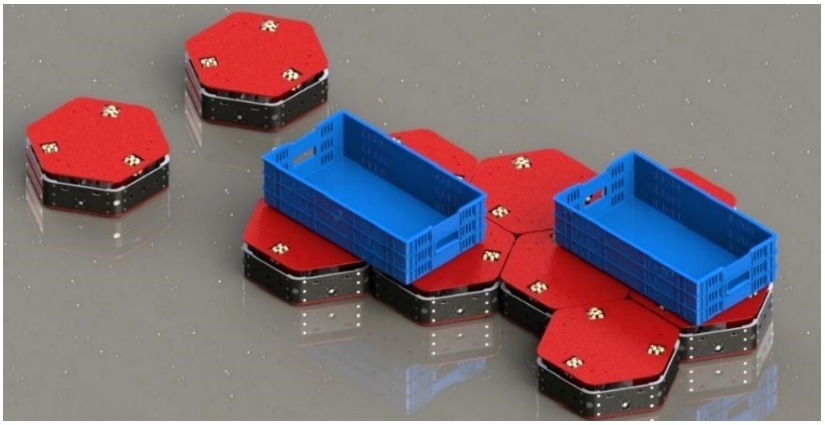


Рисунок 7 – Вариант согласованной работы нескольких транспортных роботов [13]

Для локальной системы позиционирования могут быть использованы QR-метки, SLAM (одновременная локализация и картографирование) и линии, нанесенные на пол или потолок склада готовой продукции или теплицы. За счет считывания информации с метки робот определяет текущую позицию. От скорости мобильной платформы зависит скорость обработки данных. Чем скорость больше, тем чаще необходимо обновлять данные о разметке дороги.

Основной алгоритм перемещения платформы по линии:

1. Получить размер изображения.
2. Установить частоту кадров видео.
3. Обрабатывать каждый кадр.
4. Сначала преобразовать кадр в оттенки серого, размыть его, чтобы удалить шум.
5. Установить порог с помощью OpenCV так, чтобы обнаруживались только белый и черный цвета, и удалить все шумы.
6. Инвертировать значение пикселя обрабатываемого изображения и найти контуры.
7. Чтобы удалить ненужный контур, необходимо найти его область и выбрать диапазон площади, который будет обрабатываться.
8. Найти момент контура и, используя момент, найти его центроид.
9. Устанавить диапазон центроида для прямой.

Изображение с видеокамеры обычно имеет ряд недостатков, такие как зашумленность и плохой контраст. Поэтому перед анализом изображения необходимо провести предобработку с целью повышения качества. На рисунке 8 представлен результат обработки изображения с применением данного алгоритма.

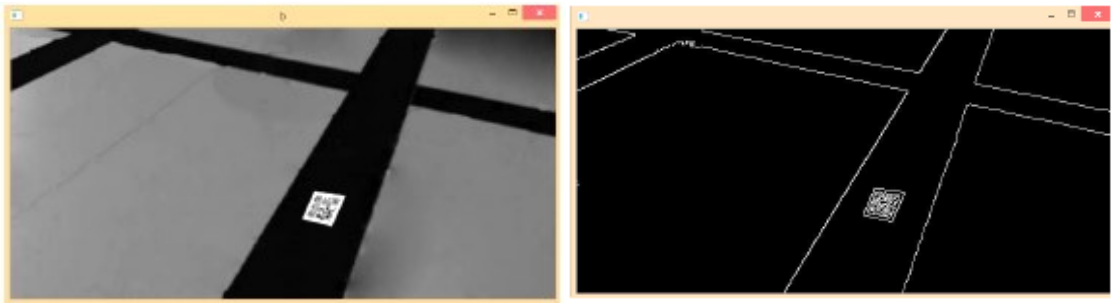


Рисунок 8 – Распознавание контуров разметки, нанесенной на пол теплицы [30]

Алгоритм обнаружения QR-кода может быть реализован при помощи библиотек pyzabr и OpenCV. Алгоритм распознавания в этом случае следующий:

1. Получение изображений с видеокамеры (покадровое считывание).
2. Перевод изображения из цветного в оттенки серого.
3. Бинаризация изображения.
4. Разметка на обособленные области (блобы).
5. Обработка получившихся областей (вычисление числовых и геометрических характеристик).
6. Фильтрация областей.
7. Поиск областей, которые по характеристикам с большей вероятностью похожи на обязательные метки позиционирования на QR-коде. Мишень – область QR-кода, образованная тремя вложенными квадратами с отношением сторон 7:5:3.

8. Выбор трех наиболее вероятных областей.

9. Анализ расположения QR-кода на изображении.

10. Выделение QR-кода из всего изображения для последующей обработки.

11. Определение необходимых преобразований.

12. Применение необходимых преобразований к QR-коду.

13. Декодирование QR-кода с помощью алгоритма, описанного в спецификации ISO/IEC 18004 «Информационная технология – Автоматическая идентификация и сбор данных – Символика штрихового кода – QR-код».

14. При удачном завершении этапа декодирования вывод на экран закодированной информации.

На рисунке 9 приведено описание полей QR-кода, на которые необходимо ориентироваться при коррекции полученных изображений.

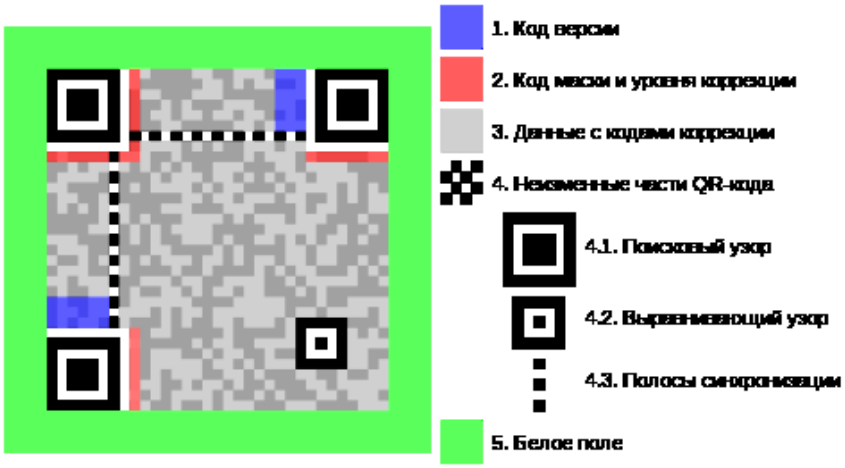


Рисунок 9 – Описание полей QR-кода [30]

Заключение. В настоящее время идёт тенденция вытеснения ручного труда, причём сельскохозяйственная область нуждается в этом не менее чем иные промышленные предприятия, автоматизация которых идёт уже давно и достигла впечатляющих результатов, чего не сказать о сельскохозяйственном секторе. Поэтому создание робототехнических систем с применением технического зрения для сельского хозяйства – задача нетривиальная и требует повышенного внимания.

Библиографический список

1. Всемирная организация здравоохранения. – Режим доступа: http:// www.who.int/ru/, свободный. – Заглавие с экрана. – Яз. рус.

2. Рыбаков, А. В. Перспективы использования мобильных роботов с системами технического зрения в сельском хозяйстве / А. В. Рыбаков, З. М. Ходарова // Симметрии: теоретический и методический аспекты : сборник научных трудов VII Международной научно-практической конференции. – 2018. – С. 39–46.

3. How Dogtooth Technologies' intelligent robots are mastering strawberry-picking. – Режим доступа: https://www.cambridgeindependent.co.uk/business/how-dogtooth-technologies-intelligent-robots-aremastering-strawberry-picking-9050467/, свободный. – Заглавие с экрана. – Яз. англ.

4. Выращивание рассады томатов в домашних условиях. – Режим доступа: https://chudoclumba.ru/vyrasivanie-rassady-tomatov-v-domasnih-usloviah/, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 22.12.2020).

5. Рыбаков, А. В. Проектирование робототехнических манипуляторов с системой компьютерного зрения для сбора томатов / А. В. Рыбаков, А. М. Лихтер, А. Б. Погожева, А. В. Михайлова, А. Б. Дусалиев // Прикаспийский журнал: управление и высокие технологии. – 2020. – № 3 (51). – С. 135–147.

6. Hongkun, Tian. Computer vision technology in agricultural automation / Hongkun, Tian, Tianhai Wang, Yadong Liu, Xi Qiao, Yanzhou Li // Information Processing in Agriculture. – 2020. – Vol. 7, iss. 1. – P. 1–19. – https://doi.org/10.1016/j.inpa.2019.09.006.

7. Perrez-Zavala, R. A pattern recognition strategy for visual grape bunch detection in vineyards / R. Perrez-Zavala, M. Torres-Torriti, F. A. Cheein, et al. // Comput. Electron. Agric. – 2018. – Vol. 151. – P. 136–149.

8. Fahmi F. Image processing analysis of geospatial uav orthophotos for palm oil plantation monitoring / F. Fahmi, D. Trianda, U. Andayani, et al. // J. Phys. Conf. Ser. – 2018. – Vol. 978. – P. 012064.

9. Sun, Y. Utilization of machine vision tomonitor the dynamic responses of rice leaf morphology and colour to nitrogen, phosphorus, and potassium deficiencies / Y. Sun, J. Gao, K. Wang, et al. // J. Spectroscopy. – 2018. – P. 1–13.

10. Zhu, Y. In-field automatic observation of wheat heading stage using computer vision / Y. Zhu, Z. Cao, H. Lu, et al. // Biosyst. Eng. – 2016. – Vol. 143. – P. 28–41.

11. Sadeghi-Tehran, P. Automated method to determine two critical growth stages of wheat: heading and flowering / P. Sadeghi-Tehran, K. Sabermanesh, N. Virlet, et al. // Front. Plant. Sci. – 2017. – № 8. – P. 252.

12. Yuan Ting, X. C. Information acquisition of cucumber fruit in greenhouse environment based on nearing infrared image / X. C. Yuan Ting // Spectrosc. Spect. Anal. – 2009. – Vol. 29 (8). – P. 2054–2058.

13. Zhang, Q. Cherry recognition in natural environment based on the vision of picking robot / Q. Zhang, S. Yu. T. Chen, et al. // IOP Conf. Ser.: Earth Environ Sci. – 2017. – Vol. 61. – Art. 012021.

14. Christopher, McCool IS. Visual detection of occluded crop: for automated harvesting / Christopher McCool IS, Feras Dayoub, Christopher Lehnert, Tristan Perez, and Ben Upcroft // International Conference on Robotics and Automation (ICRA). – 2016. – P. 2506–2512.

15. Lü, Q. Detection of hidden bruise on kiwi fruit using hyperspectral imaging and parallelepiped classification / Q. Lü and M. J. Tang // Procedia Environ. Sci. – Jan. 2012. – Vol. 12. – P. 1172–1179. – DOI: 10.1016/j.proenv.2012.01.404.

16. Zhang, S. An effective method to inspect and classify the bruising degree of apples based on the optical properties / S. Zhang, X. Wu, S. Zhang, Q. Cheng, and Z. Tan // Postharvest Biol. Technol. – May 2017. – Vol. 127. – P. 44–52. – DOI: 10.1016/j.postharvbio.2016.12.008.

17. Elien, D. Assessment of bruise volumes in apples using X-ray computed tomography / D. Elien, V. D. Mattias, and K. Janos // Postharvest Biol. Technol. – Jun. 2017. – Vol. 128. – P. 24–32. – DOI: 10.1016/j.postharvbio.2017.01.013.

18. Keresztes, J. C. Real-time pixel based early apple bruise detection using short wave infrared hyperspectral imaging in combination with calibration and glare correction techniques / J. C. Keresztes, M. Goodarzi, and W. Saeys // Food Control. – Aug. 2016. – Vol. 66, no. 1. – P. 215–226. – DOI: 10.1016/j.foodcont.2016.02.007.

19. Mochida, K. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective / K. Mochida, S. Koda, K. Inoue, et al. // GigaScience. – 2019. – № 8. – P. 1–53.

20. Vazquez-Arellano, M. 3-D imaging systems for agricultural applications / M. Vazquez-Arellano, H. W. Griepentrog, D. Reiser, et al. // A review Sensors. – 2016. – Vol. 16. – P. 1–24.

21. Mazzilli. UAV Inspection of Olive Trees for the Detection of Xylella Fastidiosa Disease Using Neural Networks," / Mazzilli et al. // 17th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA). – 2021. – P. 1–4. – DOI: 10.1109/CNNA49188.2021.9610752.

22. Jui-Cheng, Yen. A new criterion for automatic multilevel thresholding / Jui-Cheng Yen, Fu-Juay Chang, & Shyang Chang // IEEE Transactions on Image Processing. – 1995. – Vol. 4 (3). – P. 370–378. – DOI: 10.1109/83.366472.

23. Niblack, W. An Introduction to Digital Image Processing / W. Niblack. – Prentice Hall, 1986. – P. 115–116.

24. Hough, P. V. C. Method and means for recognizing complex patterns : U.S. Patent / P. V. C. Hough. – 1962, Dec. 18. – № 3, 069, 654.

25. Рыбаков, А. В. Анализ использования систем компьютерного зрения для контроля состояния рассады / А. В. Рыбаков, И. В. Михайлов, Р. К. Арыкбаев // Прикаспийский международный молодёжный научный форум агропромтехнологий и продовольственной безопасности – 2019 : сборник научных статей. – Астрахань, 23–24 апреля 2019 г. – Астрахань : Издательский дом «Астраханский университет», 2019. – С. 57–60.

26. Андреев, С. В. Построение стереоанимационных представлений на современных стереоустановках активного и пассивного типов / С. В. Андреев, А. Е. Бондарев, Т. Н. Михайлова, И. Г. Рыжова // Новые информационные технологии в автоматизированных системах. – 2018. – № 21. – С. 4–10.

27. Zuiderveld, K. Contrast Limited Adaptive Histogram Equalization / K. Zuiderveld // Graphics Gems IV. – Academic Press, 1994. – Режим доступа: http://www.docin.com/p-119164091.html, свободный. – Заглавие с экрана. – Яз. англ.

28. Нейросеть заинтересовалась сорняками. – Режим доступа: https://agronomius.livejournal.com/22747.html, свободный. – Заглавие с экрана. – Яз. рус.

29. Рыбаков, А. В. Концепция роботизированного тепличного комплекса для выращивания томатов с одним оператором / А. В. Рыбаков, Е. Ю. Степанович, И. В. Михайлов [и др.] // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 1 (53). – С. 116–126.

30. Матричный код (QR-код). – Режим доступа: http://www.russika.ru/ef.php?s=5337, свободный. – Заглавие с экрана. – Яз. рус.

References

1. *Vsemirnaya organizatsiya zdravookhraneniya* [World Health Organization]. Available at: http:// www.who.int/ru/.

2. Rybakov, A. V., Khodarova, Z. M. Perspektivy ispolzovaniya mobilnykh robotov s sistemami tekhnicheskogo zreniya v selskom hozyaystve [Prospects for the use of mobile robots with technical vision systems in agriculture]. *Simmetrii: teoreticheskiy i metodicheskiy aspekty : sbornik nauchnykh trudov VII Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Symmetries: theoretical and methodological aspects : collection of scientific papers of the VII International Scientific and Practical Conference], 2018, pp. 39–46.

3. *How Dogtooth Technologies' intelligent robots are mastering strawberry-picking.* Available at: https://www.cambridgeindependent.co.uk/business/how-dogtooth-technologies-intelligent-robots-aremastering-strawberry-picking-9050467/.

4. *Vyrashchivanie rassady tomatov v domashnikh usloviyakh* [Growing tomato seedlings at home]. Available at: https://chudoclumba.ru/vyrasivanie-rassady-tomatov-v-domasnih-usloviah/ (accessed 22.12.2020).

5. Rybakov, A. V. Likhter, A. M., Pogozheva, A. B., Mikhaylova, A. V., Dusaliev, A. B. Proektirovanie robototekhnicheskikh manipulyatorov s sistemoy kompyuternogo zreniya dlya sbora tomatov [Design of robotic

manipulators with a computer vision system for picking tomatoes]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2020, no. 3 (51), pp. 135–147.

6. Hongkun, Tian, Tianhai, Wang, Yadong, Liu, Xi, Qiao, Yanzhou, Li. Computer vision technology in agricultural automation. *Information Processing in Agriculture*, 2020, vol. 7, iss. 1, pp. 1–19. https://doi.org/10.1016/j.inpa.2019.09.006.

7. Perrez-Zavala, R., Torres-Torriti, M., Cheein, F. A., et al. A pattern recognition strategy for visual grape bunch detection in vineyards. *Comput. Electron. Agric.*, 2018, vol. 151, pp. 136–49.

8. Fahmi, F., Trianda, D., Andayani, U., et al. Image processing analysis of geospatial uav orthophotos for palm oil plantation monitoring. *J. Phys. Conf. Ser.*, 2018, vol. 978, p. 012064.

9. Sun, Y., Gao, J., Wang, K., et al. Utilization of machine vision tomonitor the dynamic responses of rice leaf morphology and colour to nitrogen, phosphorus, and potassium deficiencies. *J. Spectroscopy*, 2018, pp. 1–13.

10. Zhu, Y, Cao, Z, Lu, H., et al. In-field automatic observation of wheat heading stage using computer vision. *Biosyst. Eng.*, 2016, vol. 143, pp. 28–41.

11. Sadeghi-Tehran, P., Sabermanesh, K., Virlet, N., et al. Automated method to determine two critical growth stages of wheat: heading and flowering. *Front. Plant. Sci.*, 2017, no. 8, p. 252.

12. Yuan Ting, X. C. Information acquisition of cucumber fruit in greenhouse environment based on nearing infrared image. *Spectrosc. Spect. Anal.*, 2009, no. 29 (8), pp. 2054–2058.

13. Zhang, Q., Chen, S. Yu. T., et al. Cherry recognition in natural environment based on the vision of picking robot. *IOP Conf. Ser: Earth Environ. Sci.*, 2017, vol. 61, art. 012021.

14. Christopher, McCool IS, Feras, Dayoub, Christopher, Lehnert, Tristan, Perez, and Ben, Upcroft. Visual detection of occluded crop: for automated harvesting. *International Conference on Robotics and Automation (ICRA)*, 2016, pp. 2506–2512.

15. Lü, Q. and Tang, M. J. Detection of hidden bruise on kiwi fruit using hyperspectral imaging and parallelepiped classification. *Procedia Environ. Sci.*, Jan. 2012, vol. 12, pp. 1172–1179. DOI: 10.1016/j.proenv.2012.01.404.

16. Zhang, S., Wu, X., Zhang, S., Cheng, Q., and Tan, Z. An effective method to inspect and classify the bruising degree of apples based on the optical properties. *Postharvest Biol. Technol.*, May 2017, vol. 127, pp. 44–52. DOI: 10.1016/j.postharvbio.2016.12.008.

17. Elien, D., Mattias, V. D., and Janos, K. Assessment of bruise volumes in apples using X-ray computed tomography. *Postharvest Biol. Technol.*, Jun. 2017, vol. 128, pp. 24–32. DOI: 10.1016/j.postharvbio.2017.01.013.

18. Keresztes, J. C., Goodarzi, M., and Saeys, W. Real-time pixel based early apple bruise detection using short wave infrared hyperspectral imaging in combination with calibration and glare correction techniques. *Food Control*, Aug. 2016, vol. 66, no. 1, pp. 215–226. DOI: 10.1016/j.foodcont.2016.02.007.

19. Mochida, K., Koda, S., Inoue, K., et al. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. *GigaScience*, 2019, no. 8, pp. 1–53.

20. Vazquez-Arellano, M., Griepentrog, H. W., Reiser, D., et al. 3-D imaging systems for agricultural applications. *A Review Sensors*, 2016, vol. 16, pp. 1–24.

21. Mazzilli et al. UAV Inspection of Olive Trees for the Detection of Xylella Fastidiosa Disease Using Neural Networks. *17th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA)*, 2021, pp. 1–4. DOI: 10.1109/CNNA49188.2021.9610752.

22. Jui-Cheng, Yen, Fu-Juay, Chang, & Shyang, Chang. A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, 1995, vol. 4 (3), pp. 370–378. DOI:10.1109/83.366472.

23. Niblack, W. *An Introduction to Digital Image Processing.* Prentice Hall, 1986, pp. 115–116.

24. Hough, P. V. C. *Method and means for recognizing complex patterns : U.S. Patent.* 1962, Dec. 18, no. 3, 069, 654.

25. Rybakov, A. V., Mikhaylov I. V., Arykbaev R. K. Analiz ispolzovaniya sistem kompyuternogo zreniya dlya kontrolya sostoyaniya rassady [Analysis of the use of computer vision systems for monitoring the state of seedlings]. *Prikaspiyskiy mezhdunarodnyy molodyozhnyy nauchnyy forum agropromtekhnologiy i prodovolstvennoy bezopasnosti – 2019 : sbornik nauchnykh statey, Astrahan, 23–24 aprelya 2019 goda* [Caspian International Youth Scientific Forum of Agro-Industrial Technologies and Food Security – 2019 : collection of scientific articles. Astrakhan, April 23–24]. Astrahan, Publishing House “Astrakhan University”, 2019, pp. 57–60.

26. Andreev, S. V., Bondarev, A. E., Mikhailova, T. N., Ryzhova, I. G. Postroenie stereoanimatsionnykh predstavleniy na sovremennykh stereoustanovkakh aktivnogo I passivnogo tipov [Construction of stereoanimation representations on modern stereo installations of active and passive types]. *Novye informatsionnye tekhnologii v avtomatizorovannykh sistemakh* [New Information Technologies in Automated Systems Founders: Moscow Institute of Electronics and Mathematics of the Higher School of Economics], 2018, no. 21, pp. 4–10.

27. Zuiderveld, K. Contrast Limited Adaptive Histogram Equalization. *Graphics Gems IV*. Academic Press, 1994. Available at: http://www.docin.com/p-119164091.html.

28. *Neyroset zainteresovalas sornyakami* [The neural network is interested in weeds]. Available at: https://agronomius.livejournal.com/22747.html.

29. Rybakov A. V., Stepanovich E. Yu., Mikhailov I. V. et al. Kontseptsiya robotizirovannogo teplichnogo kompleksa dlya vyrashchivaniya tomatov s odnim operatorom [The concept of a robotic greenhouse complex for growing tomatoes with one operator]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2021, no. 1 (53), pp. 116–126.

30. *Matrichnyy kod (QR-kod)* [Matrix code (QR code)]. Available at: http://www.russika.ru/ef.php?s=5337.

ПРИБОРЫ, СИСТЕМЫ И ИЗДЕЛИЯ МЕДИЦИНСКОГО НАЗНАЧЕНИЯ

УДК 519.7

ФАКТОРНЫЙ АНАЛИЗ ВЛИЯНИЯ ПРИЗНАКОВ НА ТОЧНОСТЬ ДИАГНОСТИКИ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ ПО ДАННЫМ МИКРОВОЛНОВОЙ РАДИОТЕРМОМЕТРИИ[1]

Статья поступила в редакцию 30.12.2021, в окончательной варианте – 18.02.2022.

Гермашев Илья Васильевич, Волгоградский государственный университет, 400062, Российская Федерация, г. Волгоград, пр. Университетский, 100,
доктор технических наук, профессор, ORCID: 0000-0001-5507-8508, e-mail: germashev@ volsu.ru
Дубовская Виктория Игоревна, Волгоградский государственный университет, 400062, Российская Федерация, г. Волгоград, пр. Университетский, 100,
аспирант, ORCID: 0000-0002-4670-4682, e-mail: dubovskajav@volsu.ru
Лосев Александр Георгиевич, Волгоградский государственный университет, 400062, Российская Федерация, г. Волгоград, пр. Университетский, 100,
доктор физико-математических наук, профессор, ORCID: 0000-0002-1072-8375, e-mail: alexander.losev@volsu.ru
Попов Илларион Евгеньевич, Волгоградский государственный университет, 400062, Российская Федерация, г. Волгоград, пр. Университетский, 100,
магистрант, ORCID: 0000-0002-0997-8721, e-mail: popov.larion@volsu.ru

В работе представлена модель оценки влияния диагностических признаков онкологического заболевания на диагноз, поставленный интеллектуальной системой, в рамках конкретного метода классификации. За основу установления уровня влияния взяты факторные нагрузки признаков. Проведен сравнительный анализ факторных моделей, построенных методами главных компонент, максимального правдоподобия и наименьших квадратов.

Ключевые слова: факторный анализ, машинное обучение, консультативно-диагностические системы, рак молочной железы

FACTOR ANALYSIS OF THE INFLUENCE OF SIGNS ON THE ACCURACY OF BREAST CANCER DIAGNOSIS ACCORDING TO MICROWAVE RADIOTHERMOMETRY

The article was received by the editorial board on 30.12.2021, in the final version – 18.02.2022.

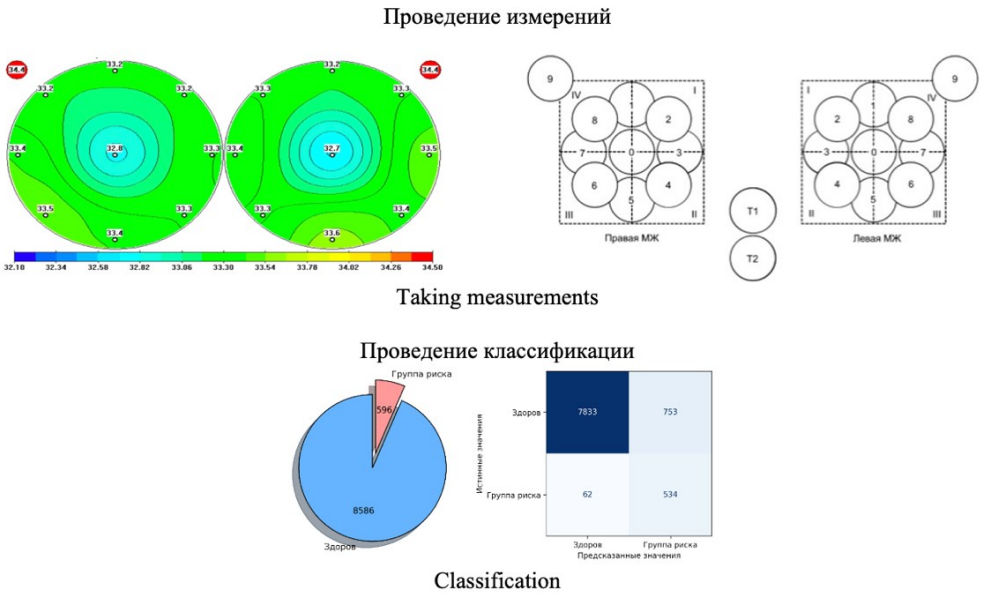
Germashev Ilya V., Volgograd State University, 100 Universitetskiy Prospect, Volgograd, 400062, Russian Federation,
Doct. Sci. (Enginering), Professor, ORCID: 0000-0001-5507-8508, e-mail: germashev@volsu.ru
Dubovskaya Victoria I., Volgograd State University, 100 Universitetskiy Prospect, Volgograd, 400062, Russian Federation,
postgraduate student, ORCID: 0000-0002-4670-4682, e-mail: dubovskajav@volsu.ru
Losev Alexander G., Volgograd State University, 100 Universitetskiy Prospect, Volgograd, 400062, Russian Federation,
Doct. Sci. (Physics and Mathematics), Professor, ORCID: 0000-0002-1072-8375, e-mail: alexander.losev@volsu.ru
Popov Illarion E., Volgograd State University, 100 Universitetskiy Prospect, Volgograd, 400062, Russian Federation,
undergraduate student, ORCID: 0000-0002-0997-8721, e-mail: popov.larion@volsu.ru

The paper presents a model for assessing the impact of diagnostic signs of cancer on the diagnosis within a specific classification method. Factor loads of features are taken as the basis for determining the level of influence. A comparative analysis of factor models constructed by the methods of principal components, maximum likelihood and least squares is carried out.

Keywords: factor analysis, machine learning, consultative and diagnostic systems, breast cancer

[1] Работа выполнена при финансовой поддержке РФФИ, проект № 19-01-00358 «Математические модели радиационных полей и анализа данных микроволновой радиотермометрии в ранней диагностике рака молочных желез».

Графическая аннотация (Graphical annotation)



Введение. Данная работа посвящена улучшению качества консультативных систем в диагностике рака молочной железы методами машинного обучения. По статистике на 2020 г., рак молочной железы является самым распространённым типом рака среди всех онкологических заболеваний у женщин и одним из наиболее распространённых среди всех типов рака [1], при этом данная тенденция носит общий характер [2, 3]. Смертность пациентов с раком молочной железы напрямую зависит от размера опухоли на момент её обнаружения [4]. Чем меньше опухоль, тем выше шанс на выздоровление при соответствующем лечении. На 2020 г. смертность от данного вида рака составляла 30 %. Таким образом, актуальной является задача ранней диагностики рака молочной железы. Одним из перспективных способов решения данной задачи нам видится применение метода микроволновой радиотермометрии (РТМ).

РТМ-метод заключается в определении поверхностных (кожных) и глубинных температур тела за счёт измерения интенсивности его электромагнитного излучения. Такие измерения становятся возможными благодаря тому, что, как и любое нагретое тело, ткани человека излучают электромагнитные колебания по всему спектру. Так, поверхностные ткани излучают колебания в инфракрасном диапазоне, а глубинные температуры – длинноволновые, т. е. микроволновые колебания [5]. Исходя из этого производятся измерения температур в различных диапазонах. Перспективность метода заключается в его неинвазивности, безопасности и возможности ранней диагностики заболеваний, в частности рака молочной железы.

Известно, что образование злокачественной опухоли в молочной железе сказывается на биохимических процессах в железе, из-за чего начинает меняться температурное поле железы [6]. Так, например, одной из характеристических особенностей заболевания является повышенная температура в некоторой области, рядом с опухолью. Благодаря температурным аномалиям такого рода становится возможной диагностика по данным, полученным РТМ-методом. При этом возможной становится как диагностика врачом-диагностом по термограммам молочных желёз, так и автоматическая диагностика методами машинного обучения. По термограммам врач-диагност оценивает состояние температур пациента, наличие температурных аномалий и характерных паттернов для рака молочной железы. На рисунке 1 приведён пример термограммы кожных температур здорового пациента, по которому наблюдаются следующие закономерности: у пациента отсутствуют области с повышенной температурой, а также температурные поля обеих молочных желёз похожи друг на друга, что характерно для здоровых пациентов. Методами же машинного обучения исследуется возможность диагностики по числовым данным температур.

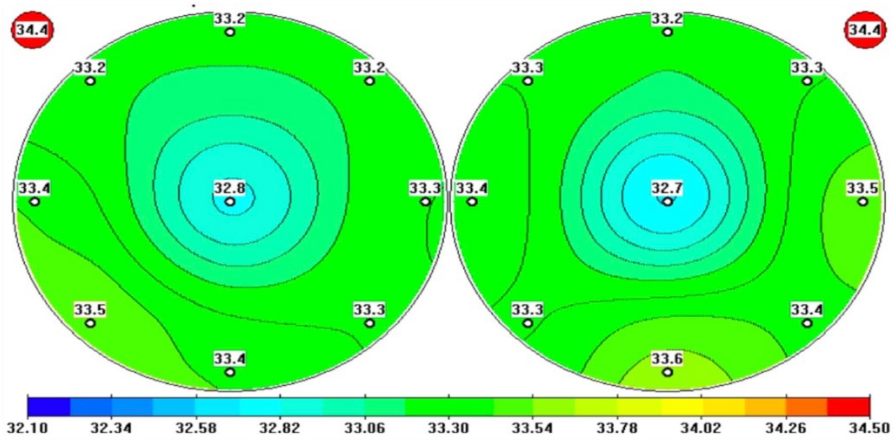


Рисунок 1 – Пример термограммы здорового пациента

Процесс построения интеллектуальных систем с использованием методов машинного обучения по результатам обследования сводится к четырём основным этапам [7]. Сначала строится дескриптивная модель, которая включает в себя гипотезы, характеризующие различные диагностические классы. В случае рака молочной железы одной из актуальных гипотез является наличие у пациентов с опухолью области с повышенной температурой. Затем строится математическая модель, которая заключается в математическом описании составленных гипотез. По математической модели проектируется признаковое пространство, используемое в алгоритмах классификации. Далее происходит построение алгоритмов машинного обучения, которое заключается в настройке и обучении алгоритмов классификации. Полученные алгоритмы классификации показывают высокую точность диагностики пациентов как с раком молочной железы [9], так и пациентов с подозрением на рак (группа риска) [8]. На последнем этапе используется механизм логического вывода результата.

Рассматриваемая концепция улучшения консультативных систем заключается в нескольких направлениях:

1) построение обоснования полученного диагноза в результате проведенной классификации на понятном специалисту языке;

2) повышение точности и адекватности моделей, лежащих в основе алгоритмов диагностики.

Вышеуказанные задачи могут быть решены в случае определения уровня влияния признаков на правильность/ошибочность поставленного классификатором диагноза. Для этого в ходе данной работы предполагается использование факторного анализа, в основе которого лежит выявление непосредственно ненаблюдаемых и неизмеряемых характеристик (общих факторов) посредством анализа вариаций наблюдаемых переменных и их корреляций.

Материалы и методы. Изначальная база данных представляет собой результаты обследований, проведённых медицинскими учреждениями. В процессе обследования температуры измерялись по схеме, показанной на рисунке 2.

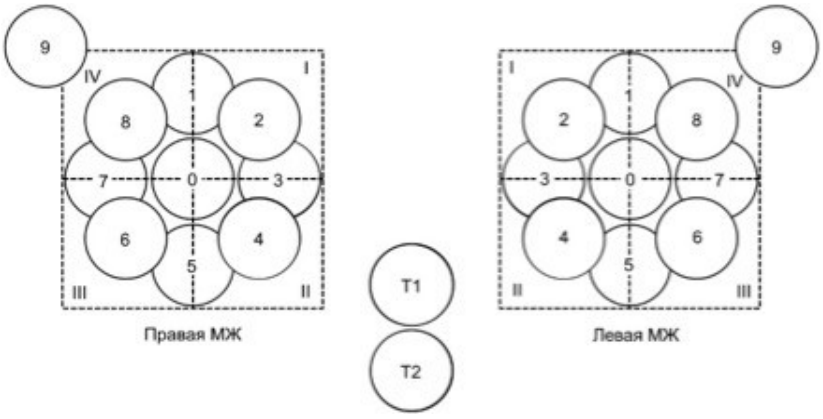


Рисунок 2 – Схема температурных измерений

Помимо температур в базу данных также записывается анамнез, который на данном этапе в алгоритмах классификации не используется, а также диагноз по каждой молочной железе пациента. Опишем данные подробнее:

- $ir_r^i = \{t^i_{0,\text{ir},r}, .., t^i_{9,\text{ir},r}\}$ – множество кожных температур правой молочной железы. Здесь и далее первый нижний индекс обозначает область, в которой измеряется температура. Значение индекса соответствует номеру области на рисунке 2. i – номер обследуемого пациента;
- $\text{mw}_\text{r}^\text{i} = \{\text{t}^\text{i}_{0,\text{mw,r}}, .., \text{t}^\text{i}_{9,\text{mw,r}}\}$ – множество глубинных температур правой молочной железы;
- $ir_l^i = \{t^i_{0,\text{ir},l}, .., t^i_{9,\text{ir},l}\}$ – множество кожных температур левой молочной железы пациента;
- $mw_l^i = \{t^i_{0,\text{mw},l}, .., t^i_{9,\text{mw},l}\}$ – множество глубинных температур левой молочной железы пациента;
- $ir_a^i = \{t^i_{0,\text{ir,a}}, t^i_{1,\text{ir,a}}\}$ – множество кожных температур опорной области (соответствует точкам T1 и T2 на рисунке 1);
- $mw_a^i = \{t^i_{0,mw,\text{a}}, t^i_{1,\text{mw,a}}\}$ – множество глубинных температур опорной области пациента;
- y_r – диагноз правой молочной железы;
- y_l – диагноз левой молочной железы.

Так как диагностика производится отдельно по каждой из молочных желёз, база данных требует некоторых изменений. Она преобразовывалась таким образом, что каждая запись в ней соответствует описанию одной молочной железы, а не пациенту с обеими железами. Для этого каждая запись дублировалась, и в них вносились следующие изменения:

- оригинальная запись соответствует правой молочной железе;
- дублирующая запись соответствует левой молочной железе;
- в каждой записи остаётся только диагноз соответствующей молочный железы;
- индексы $'r'$ и $'l'$ заменяются на $'c'$ и $'p'$, где $'c'$ означает, что температуры принадлежат молочной железе записи, $'p'$ – температуры принадлежат парной молочной железе пациента.

Таким образом была получена база данных по молочным железам. В ней содержится 8586 записей о здоровых железах и 596 записей о железах, входящих в группу риска (рис. 3). В данную группу входят записи с температурными аномалиями, характерными для молочных желез со злокачественной опухолью.

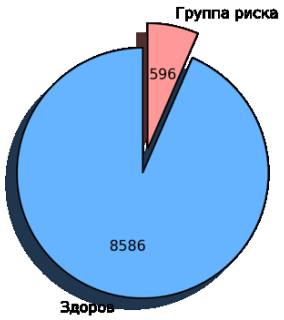


Рисунок 3 – Соотношение классов в базе данных

По полученным данным обследований пациентов специалистами была составлена следующая дескриптивная модель, характеризующая отличительные особенности пациентов из группы риска:

- температурные аномалии, возникающие из-за опухоли, делают значительными различия температурных полей парных молочных желез;
- повышенная разность температур сосков;
- неравномерное распределение температур в молочной железе, наличие «горячих» областей;
- повышенная разность глубинных и кожных температур молочной железы;
- и некоторые другие;

По составленной дескриптивной модели строилось её математическое описание и признаковое пространство. Для этого использовался набор норм и полунорм некоторых функциональных пространств. В результате были построены различные признаки, описывающие обозначенные выше особенности пациентов из группы риска [8].

Для проведения факторного анализа, с целью установления уровня влияния признаков на поставленный диагноз, необходимо было проклассифицировать все записи из базы данных одним из алгоритмов машинного обучения. В качестве такого алгоритма была выбрана логистическая

регрессия. Для того чтобы проклассифицировать каждую запись, решено было использовать метод кросс-валидации. Метод заключается в последовательном обучении и тестировании алгоритма классификации на разных подвыборках, перемешивающихся друг с другом в обучающей выборке. Для этого вся база данных делилась на 10 равных и непересекающихся частей (максимальное различие в их размере – две записи). В каждую часть входит равное количество записей каждого класса. Далее 9 из 10 частей помещались в выборку, по которой обучалась логистическая регрессия. На оставшейся части, не вошедшей в обучающую выборку, проводилась классификация. Результаты классификации сохранялись в итоговую базу данных H. Таким образом, в ней, помимо реального класса, хранится и класс, поставленный логистической регрессией. После того как все записи из рассматриваемой части были проклассифицированы, алгоритм переобучался на новой обучающей выборке. Новая выборка составлялась таким образом, что одна из частей предыдущей выборки менялась местами с классифицируемой выборкой. Процесс классификации повторялся 10 раз, пока все записи из базы данных не были проклассифицированы. На рисунке 4 приведены результаты классификации в виде матрицы ошибок. Таким образом, логистическая регрессия проклассифицировала верно 90 % здоровых молочных желёз и 88 % желёз из группы риска.

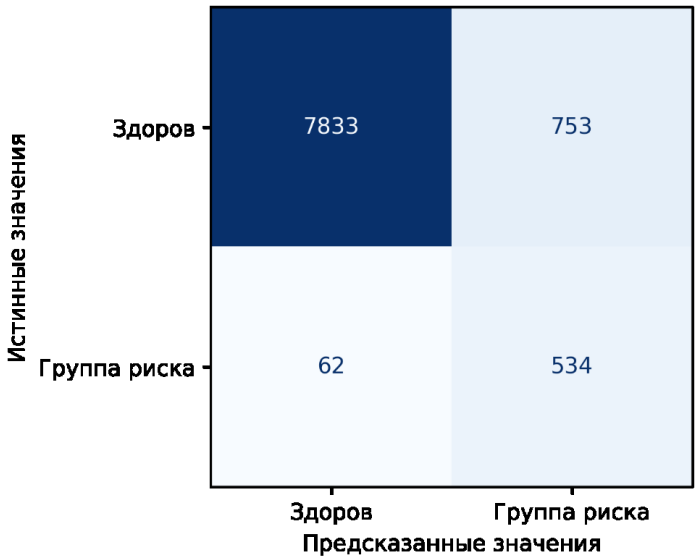


Рисунок 4 – Результаты классификации

На основе полученного результата классификации база H была разделена на две выборки:

1) H^0, состоящая из 815 элементов, представляющая собой данные молочных желез, для которых алгоритм сделал ошибочную классификацию;

2) H^1, состоящая из 8367 элементов, представляющая собой данные молочных желез, для которых алгоритм классификации определил верный диагноз.

Для исследования скрытой структуры вышеуказанных выборок без предположения о числе факторов и их нагрузках использовался разведочный факторный анализ.

Поставленная задача была реализована в трех основных этапах:

- подготовка соответствующей матрицы корреляций;
- выделение первоначальных факторов;
- получение окончательного решения с помощью вращения.

Для получения базовой модели использовался метод главных компонент. При детальном рассмотрении анализа главных компонент и, например, метода главных факторов, который является методом факторного анализа, можно отметить их сходство. Направления главных осей определяются путем нахождения собственных чисел и векторов корреляционной матрицы.

Уравнение нахождения собственных чисел и векторов в матричной записи имеет вид:

$$RV = \lambda V, \quad (1)$$

где R – исходная матрица корреляций; V – собственный вектор; λ – собственное число.

Решение основано на более простой форме в виде детерминанта матрицы:

$$Det(R - I\lambda) = 0. \quad (2)$$

При проведении факторного анализа используется характеристическое уравнение (2), как и в анализе главных компонент, но в методе главных факторов в качестве R выступает редуцированная

корреляционная матрица, у которой на главной диагонали вместо единиц расположены оценки общности. Для вычисления последних используют, например, квадрат множественного коэффициента корреляции между данной переменной и совокупностью остальных переменных, или максимальный по абсолютной величине коэффициент корреляции в строке матрицы. После расположения оценок общностей выделяются факторы аналогичным образом [10].

Однако при использовании метода главных компонент происходит поиск факторов, которые объясняют всю изменчивость исходного набора признаков. Далее факторы, которые объясняют наименьшую долю изменчивости, отбрасываются, так как считаются несущественными. При этом факторы представляют собой линейные комбинации исходных переменных. Таким образом, предположений о структуре изучаемых признаков нет.

В основе же факторного анализа лежит гипотеза о том, что изменчивость – результат влияния определенного числа факторов-причин. При этом учитывается и наличие специфических факторов, что позволяет предотвратить потерю информации.

Поэтому с целью приближения модели к реальным данным предлагается использовать такие методы факторного анализа, как метод максимального правдоподобия и наименьших квадратов.

Результаты исследований. Для проведения факторного анализа было отобрано 62 диагностических признака. Основанием для выбора данных признаков являлись значения их информативности. В данном контексте термин информативность рассматривается как количественный параметр, определяющий, насколько хорошо закономерность описывает различия между искомой и отделяемой группами.

Рассмотрим исходную матрицу корреляций r_{ij} между переменными-признаками для выборки H^0, фрагмент, которой представлен в таблице 1.

Таблица 1 – Фрагмент матрицы корреляций Пирсона

	MG001	MG002	MG003	MG004	MG005	MG006	MG007	MG008
MG001	1	0,977	0,797	0,669	0,036	–0,092	0,350	–0,077
MG002	0,977	1	0,894	0,753	0,036	–0,095	0,394	–0,082
MG003	0,797	0,894	1	0,806	0,034	–0,094	0,430	–0,090
MG004	0,669	0,753	0,806	1	0,025	–0,061	0,477	–0,057
MG005	0,036	0,036	0,034	0,025	1	–0,576	0,308	–0,423
MG006	–0,092	–0,095	–0,094	–0,061	–0,576	1	0,433	0,738
MG007	0,350	0,394	0,430	0,477	–0,308	0,433	1	0,318
MG008	–0,077	–0,082	–0,090	–0,057	–0,423	0,738	0,318	1

Данная матрица не является положительно определенной, что создает затруднения при проведении факторного анализа предпочтительными методами. Одна из возможных причин – определитель матрицы стремится к 0 либо является отрицательным, вследствие линейной зависимости между переменными. Несмотря на то, что факторный анализ нацелен на поиск подобных зависимостей, в случае когда такие зависимости перестают быть вероятностными, становятся жестко детерминированными, многомерные методы анализа могут давать сбой.

Как можно заметить, согласно таблице Чеддока, сила корреляционной связи между многими переменными характеризуется как высокая (значения коэффициентов в таблице 1 больше 0,7).

Для устранения вышеуказанной проблемы на этапе подготовки исходной матрицы предлагается оставить те признаки, для которых $r_{ij} \leq 0{,}7$.

Таким образом, для проведения факторного анализа было отобрано 25 признаков. Фрагмент списка диагностических признаков приведен в таблице 2.

Таблица 2 – Фрагмент списка высокоинформативных диагностических признаков

Название признака	Математическое описание признака	Название признака	Математическое описание признака
MG001	$\left\|T_c^{i,mw} - T_p^{i,mw}\right\|_1$	MG025	$max_{t\in T_c^{i,g}} t - min_{t\in T_p^{i,g}} t$
MG005	$s\left(T_c^{i,mw}\right) - s\left(T_p^{i,mw}\right)$	MG026	$T_{0,c}^{i,g} - T_{0,p}^{i,g}$
MG006	$\overline{T_c^{i,mw}} - \overline{T_p^{i,mw}}$	MG027	$T_{9,c}^{i,g} - T_{9,p}^{i,g}$
MG007	$max_{t\in T_c^{i,mw}} \mathrm{t} - \min_{t\in T_p^{i,mw}} \mathrm{t}$	MG032	$\min_{t\in T_a^{i,g}} \lvert t\rvert$
MG008	$T_{0,c}^{i,mw} - T_{0,p}^{i,mw}$	MG036	$\max_{t\in T^{i,mw}} \left\lvert \overline{T^{i,mw}} - t\right\rvert$
MG010	$\left\|T_c^{i,ir} - T_p^{i,ir}\right\|_1$	MG048	$\max_{t\in T^{i,ir}\setminus T_0^{i,ir}} \lvert T_0^{i,ir} - t\rvert$
MG014	$s\left(T_c^{i,ir}\right) - s\left(T_p^{i,ir}\right)$	MG053	$\max_{t\in T^{i,ir}} \lvert T_{2,a}^{i,ir} - t\rvert$
MG019	$\left\|T_c^{i,g} - T_p^{i,g}\right\|_1$	MG054	$max_{t\in T^{i,g}} \left\lvert \overline{T^{i,g}} - t\right\rvert$
MG024	$\overline{T_c^{i,g}} - \overline{T_p^{i,g}}$	MG061	$\max\limits_{t\in T^{i,g}} \lvert T_{1,a}^{i,g} - t\rvert$
MG062	$\max\limits_{t\in T^{i,g}} \lvert T_{2,a}^{i,g} - t\rvert$		

Метод максимального правдоподобия и наименьших квадратов основан на гипотезе, что наблюдаемые данные представляют собой выборку, соответствующую *k*-факторной модели. Поэтому для определения числа факторов за основу взята методика, в которой число факторов определено в ходе использования метода главных компонент при помощи критерия Кайзера и Кеттелла. Предполагается, что число факторов определяется числом компонент, собственные значения которых больше 1 (табл. 3).

Таблица 3 – Объясненная совокупная дисперсия

Компонент	Извлечение суммы квадратов нагрузок			Ротация суммы квадратов нагрузок		
	Всего	% дисперсии	Суммарный %	Всего	% дисперсии	Суммарный %
1	4,911	21,351	21,351	3,940	17,131	17,131
2	2,436	10,593	31,944	2,242	9,748	26,879
3	2,252	9,792	41,736	2,217	9,638	36,517
4	2,097	9,119	50,855	2,099	9,127	45,644
5	1,475	6,414	57,269	2,000	8,698	54,342
6	1,458	6,338	63,608	1,649	7,169	61,511
7	1,273	5,534	69,141	1,497	6,507	68,018
8	1,106	4,809	73,950	1,364	5,932	73,950

Таким образом, на основании данных из таблицы 3, число факторов *k* принимаем равным 8, и эти факторы описывают 73,9 % совокупной дисперсии признаков.

Далее был проведен факторный анализ методом максимального правдоподобия, в результате которого было извлечено 8 факторов за 15 итераций. Вращение по методу варимакс с нормализацией Кайзера сошлось за 6 итераций. Полученная матрица факторных нагрузок c^0 приведена в таблице 4.

Таблица 4 – Факторная матрица для выборки H^0 после вращения

Признак	1	2	3	4	5	6	7	8
MG065	0,845	0,021	0,174	–0,060	0,182	–0,031	0,020	0,216
MG061	0,842	0,074	–0,086	–0,051	0,120	0,083	–0,027	–0,106
MG025	0,834	0,014	–0,017	0,010	0,128	0,015	0,293	–0,090
MG064	0,741	0,034	0,071	–0,074	0,138	–0,013	0,021	0,220
MG019	0,553	–0,025	0,127	0,346	–0,009	–0,112	–0,043	–0,019
MG048	0,542	0,342	0,090	0,138	0,034	0,539	0,054	–0,078
MG005	0,015	0,920	0,016	0,020	0,001	–0,031	–0,022	0,033
MG014	0,122	0,818	0,012	–0,018	–0,002	0,063	0,012	–0,024
MG008	0,036	–0,450	0,008	–0,086	–0,282	0,214	–0,053	0,015
MG063	–0,233	0,031	–0,961	0,037	–0,134	–0,001	–0,025	–0,030

Аналогичным образом была получена матрица факторных нагрузок с использованием метода наименьших квадратов (извлечено 8 факторов, потребовалось 17 итераций, метод вращения – варимакс с нормализацией Кайзера, вращение сошлось за 7 итераций).

Таблица 5 – Факторная матрица для выборки H^0 после вращения (метод наименьших квадратов)

Признак	1	2	3	4	5	6	7	8
MG065	0,856	0,020	0,177	–0,081	0,177	–0,036	0,022	0,169
MG025	0,834	0,014	–0,019	0,009	0,163	0,063	0,288	–0,082
MG061	0,819	0,089	–0,100	–0,047	0,139	0,131	–0,025	–0,046
MG064	0,725	0,043	0,058	–0,079	0,131	0,002	0,017	0,210
MG019	0,566	–0,026	0,130	0,245	–0,038	–0,127	–0,035	–0,048
MG048	0,527	0,358	0,096	0,141	0,046	0,512	0,012	0,005
MG005	0,017	0,935	0,020	0,009	–0,006	–0,052	–0,030	0,013
MG014	0,115	0,791	–0,018	–0,039	0,049	0,076	–0,025	–0,032
MG008	0,011	–0,428	–0,050	–0,109	–0,129	0,193	–0,065	–0,017
MG063	–0,252	0,029	–0,955	0,055	–0,117	0,049	0,012	–0,028

Полученные факторные нагрузки для выборки H^0 можно интерпретировать как уровень влияния признака на ошибочность результата классификации в рамках конкретного метода машинного обучения.

В процессе проведения факторного анализа также рассчитывались и анализировались следующие показатели:

1) критерий сферичности Бартлетта – так как уровень значимости для каждого из примененного метода меньше 0,05, то можно сделать вывод о приемлемости проведения факторного анализа;

2) KMO (мера адекватности выборки Кайзера – Майера – Олкина) – значения для методов примерно одинаковы и ≈ 0.65, таким образом можно говорить об адекватности проведенного анализа и его применимости к данной выборке.

Таким образом, и один, и другой метод могут быть применены для решения поставленной задачи. Полученные факторные модели являются адекватными и применимыми на практике.

Для проведения дальнейшего анализа использовались факторные нагрузки, полученные с помощью метода максимального правдоподобия. Важно отметить, что переменные, имеющие нулевой вклад в ошибочность решения ($c_i^0 = 0$), одни и те же в результатах, полученных обоими методами. К ним относятся, например, асимметрии внутренних температур в манхэттенской метрике.

Такой же алгоритм получения факторных нагрузок c^1 был использован и для выборки H^1 (табл. 6).

Таблица 6 – Факторная матрица после вращения (выборка H^1)

Признак	1	2	3	4	5	6	7	8
MG065	0,832	0,120	0,040	0,072	–0,024	–0,024	0,127	0,004
MG061	0,777	0,129	0,076	0,156	0,303	–0,013	–0,011	–0,019
MG025	0,770	0,170	0,072	–0,002	0,144	–0,030	0,003	–0,052
MG064	0,747	0,094	0,029	0,085	0,056	–0,022	0,033	0,143
MG024	0,116	–0,031	0,002	–0,017	–0,020	0,019	–0,017	0,007
MG010	0,158	0,982	0,089	–0,011	–0,026	0,014	–0,023	–0,008
MG001	–0,025	0,768	0,266	–0,007	0,029	0,028	–0,064	0,085
MG019	0,413	0,512	–0,010	–0,017	–0,041	–0,011	0,038	–0,047

Полученные факторные нагрузки для выборки H^1 можно интерпретировать как уровень влияния признака на правильность поставленного диагноза.

Значения c^0 и c^1 позволяют оценить уровень доверия к признаку при постановке диагноза.

На их основе стало возможным разделение признаков на три различные группы:

1) признаки, имеющие высокое значение c_{ij}^1;

2) признаки, имеющие высокое значение c_{ij}^0;

3) признаки, имеющие высокое значение c_{ij}^0 и c_{ij}^1.

Выводы. Основной вклад в установление правильного диагноза имеют признаки первой группы, для которых уровень влияния $c_{ij}^0 = 0$. Таким образом, данные признаки наиболее точно отмечают тепловые изменения в органе. В данной группе преобладают признаки, касающиеся внутренних температур. Следовательно, в обосновании результатов их необходимо поместить в начале иерархии признаков. К таковым признакам относятся: асимметрия внутренних температур по области в манхэттенской метрике, в сосках, в аксиллярной области, асимметрия внутренних градиентов в тех же точках. Проведя классификацию молочных желез, основанную только на этих признаках, получим результат (специфичность 0,88, чувствительность 0,85, эффективность 0,86), который является сопоставимым с представленным в разделе 2, при котором использовались все 62 признака.

Превалирующее количество признаков, для которых уровень влияния $c_{ij}^1 = 0$, относится к группе функционалов, моделирующих поведение температурных полей в ИК-диапазоне. Например, среднеквадратическое отклонение и колебание температур молочной железы, отклонение температуры соска, аксиллярной области, отклонение кожных температур в опорных точках. Следовательно, логично поместить данные признаки ниже по иерархии, так как они могут носить не основной, а уточняющий характер.

Для тех признаков, у которых уровень влияния на правильность и ошибочность одинаково высок, предлагается использование нечеткого вывода Мамдани для получения более однозначного решения.

Использование методов максимального правдоподобия и наименьших квадратов является более предпочтительным для данного исследования, так как увеличивается адекватность модели и ее соответствие реальным данным. При этом более точно удается идентифицировать признаки, имеющие высокий уровень влияния на правильность результата. Если сравнивать предлагаемые здесь методы максимального правдоподобия и наименьших квадратов с методом главных компонент [11], то можно сделать следующие выводы. Количество признаков в первой группе признаков, которые имеют сильное влияние на правильность поставленного диагноза, увеличилось на 6 признаков, во второй группе признаков, имеющих сильное влияние на ошибочность результата классификации, количество осталось неизменным и равно 9.

Библиографический список

1. Sung, H. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries / H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray // CA Cancer J. Clin. – 2021. – Vol. 71. – P. 209–249. – DOI: 10.3322/caac.21660.

2. Bray, F. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries / F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal // CA: A Cancer Journal for Clinicians. – 2018. – Vol. 68. – P. 394–424. – DOI: 10.3322/caac.21492.

3. Siegel, R. L. Cancer statistics / R. L. Siegel, K. D. Miller and A. Jemal // CA: A Cancer Journal for Clinicians. – 2016. – Vol. 66. – P. 7–30. – DOI: 10.3322/caac.21332.

4. Michaelson, J. S. Predicting the survival of patients with breast carcinoma using tumor size / J. S. Michaelson, M. Silverstein, J. Wyatt, G. Weber, R. Moore, E. Halpern, D. B. Kopans, K. Hughes // Cancer. – 2002, Aug 15. – Vol. 95 (4). – P. 713–723. – DOI: 10.1002/cncr.10742.

5. Vesnin, S. Modern Microwave Thermometry for Breast Cancer / S. Vesnin, A. K. Turnbull, J. M. Dixon, I. Goryanin // Journal of Molecular Imaging & Dymanics. – 2017. – Vol. 7, iss. 2. – DOI: 10.4172/2155-9937. 1000136.

6. Chanmugam, A. Thermal analysis of cancerous breast model / A. Chanmugam, R. Hatwar, C. Herman // Int. Mech. Eng. Congress Expo. – 2012. – P. 134–143. – DOI: 10.1115/IMECE2012-88244.

7. Losev, A. G. Intellectual analysis of microwave radiothermometry data in the diagnosis of breast cancer / A. G. Losev, V. V. Levshinsky // Mathematical physics and computer modeling. – 2017. – Vol. 20, № 5. – P. 49–62. DOI: 0.15688/mpcm.jvolsu.2017.5.6

8. Levshinskii, V. V. Mathematical models for analyzing and interpreting microwave radiometry data in medical diagnosis / V. V. Levshinskii // Journal of Computational and Engineering Mathematics. – 2021. – Vol. 8, № 1. – P. 3–14. – DOI: 10.14529/jcem210101.

9. Losev, A. G. Neural Networks in Diagnosis of Breast Cancer/ A. G. Losev, D. A. Medevedev, A. V. Svetlov // “Smart Technologies” for Society, State and Economy. ISC 2020. Lecture Notes in Networks and Systems. – Springer, Cham., 2021. – Vol. 155. – DOI: 10.1007/978-3-030-59126-7_25.

10. Окунь, Ян. Факторный анализ / Ян. Окунь ; пер. с польск. Г. З. Давидовича ; науч. ред. В. М. Жуковская. – Москва : Статистика, 1974. – 199 с.

11. Гермашев, И. В. Модель иерархии признаков в диагностике рака молочной железы по данным микроволновой радиотермометрии / И. В. Гермашев, В. И. Дубовская, А. Г. Лосев // Математические методы в технологиях и технике. – 2021. – № 6. – С. 78–83.

References

1. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, 2021, vol. 71, pp. 209–249. DOI: 10.3322/caac.21660.

2. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. and Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 2018, vol. 68, pp. 394–424. DOI: 10.3322/caac.21492.

3. Siegel, R. L., Miller, K. D. and Jemal, A. Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians*, 2016, vol. 66, pp. 7–30. DOI: 10.3322/caac.21332.

4. Michaelson, J. S., Silverstein, M., Wyatt, J., Weber, G., Moore, R., Halpern, E., Kopans, D. B., Hughes, K. Predicting the survival of patients with breast carcinoma using tumor size. *Cancer*, 2002, Aug 15, vol. 95 (4), pp. 713–23. DOI: 10.1002/cncr.10742.

5. Vesnin, S., Turnbull, A. K., Dixon, J. M., Goryanin, I. Modern Microwave Thermometry for Breast Cancer. *Journal of Molecular Imaging & Dymanics*, 2017, vol. 7, iss. 2. DOI: 10.4172/2155-9937.1000136.

6. Chanmugam, A., Hatwar, R., Herman, C. Thermal analysis of cancerous breast model. *Int. Mech. Eng. Congress Expo*, 2012, pp. 134–143. DOI: 10.1115/IMECE2012-88244.

7. Losev, A. G., Levshinsky, V. V. Intellectual analysis of microwave radiothermometry data in the diagnosis of breast cancer. *Mathematical physics and computer modeling*, 2017, vol. 20, no. 5, pp. 49–62. DOI: 0.15688/mpcm. jvolsu.2017.5.6.

8. Levshinskii, V. V. Mathematical models for analyzing and interpreting microwave radiometry data in medical diagnosis. *Journal of Computational and Engineering Mathematics*, 2021, vol. 8, no. 1, pp. 3–14. DOI: 10.14529/jcem210101.

9. Losev, A. G., Medevedev, D. A., Svetlov, A. V. Neural Networks in Diagnosis of Breast Cancer. *Smart Technologies"for Society, State and Economy. ISC 2020. Lecture Notes in Networks and Systems*. Springer, Cham., 2021, vol. 155. DOI: 10.1007/978-3-030-59126-7_25.

10. Okun, Yan. *Faktornyy analiz* [Factor analysis]. Moscow, Statistika Publ., 1974. 199 p.

11. Germashev, I. V., Dubovskaya, V. I., Losev, A. G. Model ierarkhii priznakov v diagnostike raka molochnoy zhelezy po dannym mikrovolnovoy radiotermometrii [A model of the hierarchy of signs in the diagnosis of breast cancer according to microwave radiothermometry]. *Matematicheskiye metody v tekhnologiyakh i tekhnike* [Mathematical methods in technology and engineering], 2021, no. 6, pp. 78–83.

ИНФОРМАЦИЯ О ПРЕДСТОЯЩИХ МЕРОПРИЯТИЯХ

30 мая – 3 июня 2022 г., Ярославский государственный технический университет, Ярославский государственный университет, Академия Пастухова, г. Ярославль, Россия

Международная научная мультиконференция
«КИБЕРФИЗИЧЕСКИЕ СИСТЕМЫ: ПРОЕКТИРОВАНИЕ И МОДЕЛИРОВАНИЕ»
«CYBER-PHYSICAL SYSTEMS DESIGN AND MODELLING» (CyberPhy – 2022)
(Scopus, Springer)

Секции:

1. Проектирование киберфизических систем (Design of cyber-physical systems).
2. Моделирование киберфизических систем и цифровых двойников (Cyber-physical systems and digital twins).
3. Интеллектуальное управление киберфизическими системами (Intelligent control of cyber-physical systems).
4. Интернет вещей для промышленных киберфизических систем (IoT for industrial cyber-physical systems).
5. Промышленная кибербезопасность (Industrial cybersecurity).
6. Инженерное образование для разработки киберфизических систем (Engineering education for cyber-physical systems development).
7. Society 5.0: Cyberspace for advanced human-centered society (Общество 5.0: киберпространство для развитого общества, ориентированного на человека).

XXXIV Международная научная конференция
«МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ТЕХНИКЕ И ТЕХНОЛОГИЯХ – ММТТ-35»
(elibrary, DOI)

Секции:

1. Качественные и численные методы исследования дифференциальных и интегральных уравнений.
2. Оптимизация, автоматизация и оптимальное управление технологическими процессами.
3. Математическое моделирование технологических и социальных процессов.
4. Математическое моделирование и оптимизация в задачах САПР, аддитивных технологий, цифрового производства.
5. Математические методы в задачах радиотехники, радиоэлектроники и телекоммуникаций, геоинформатики, авионики и космонавтики.
6. Математические методы и интеллектуальные системы в робототехнике и мехатронике.
7. Математические методы в медицине, биотехнологии и экологии.
8. Математические методы в экономике и гуманитарных науках.
9. Информационные и интеллектуальные технологии в технике и образовании.
10. Математические и инструментальные методы технологий Индустрии 4.0.
11. Обсуждение квалификационных работ.
12. Круглые столы.

Подача заявок на участие с 25 декабря 2021 г.
Подробная информация о конференции и условиях участия в ней размещена на сайте
http://mmtt.sstu.ru/

ПРАВИЛА ДЛЯ АВТОРОВ

1. В журнале публикуются материалы на английском и русском языках по тематике, соответствующей утвержденным для журнала отраслям наук, группам специальностей.

2. В список соавторов работ включаются только те лица, которые внесли творческий вклад в подготовку представленных материалов. Лицам, оказавшим только техническую помощь, можно выразить благодарность в конце статьи. Один человек может быть автором (соавтором) не более чем двух статей в одном номере журнала, причем единственным автором он может быть только в одной статье.

3. Объем публикаций для научных статей должен быть не менее 8 страниц, а количество источников в библиографическом списке (списке литературы) – не менее 10 позиций.

4. Содержание каждой статьи должно включать следующие элементы: УДК; название статьи; сведения об авторах, включая их место работы, должность, адрес электронной почты; аннотацию объемом от 100 до 250 слов, ключевые слова (от 9 до 13); графическую аннотацию, отражающую содержание статьи; название статьи, сведения об авторах, аннотацию и ключевые слова на английском языке (для англоязычных статей – на русском языке); введение – оно должно заканчиваться формулировкой цели работы в явной форме; собственно текст статьи – очень желательна его сегментация на разделы, имеющие содержательные заголовки; выводы или заключение (должны соответствовать формулировке цели статьи).

5. Для русскоязычных статей приводится два библиографических списка: на языке оригинала статьи; список с транслитерацией русскоязычных источников на латиницу и (дополнительно) приведением в квадратных скобках переводов названий статей и названий источников на английский язык.

В «русскоязычном» библиографическом списке (списке литературы) порядок следования источников – по алфавиту фамилий авторов (сначала русскоязычные источники, потом иноязычные). На все источники, включенные в библиографический список, должны быть даны ссылки в тексте статьи в квадратных скобках. При необходимости авторы могут указывать номера страниц в источниках, на которые даются ссылки. Приветствуются ссылки на иноязычные источники, а также на материалы, опубликованные ранее в журнале «Прикаспийский журнал: управление и высокие технологии». Однако в последнем случае количество таких ссылок не должно превышать 20 % от общего количества источников, включенных в библиографический список. Для источников, имеющих DOI, целесообразно его указывать. При ссылках на статьи, опубликованные в журнале «Прикаспийский журнал: управление и высокие технологии», целесообразно в конце библиографического описания источника в круглых скобках указывать гиперссылку, указывающую на место размещения статьи на страничке сайта Астраханского государственного университета.

Ссылки в библиографическом списке на материалы, размещенные в интернете, допускаются при соблюдении следующих условий: если у материала, на который дается ссылка, имеется автор и/или название, то они должны быть указаны для этого источника; должен быть приведен полный маршрут доступа к источнику в интернете; должна быть указана дата обращения (доступа) к источнику.

Ограничения по списку литературы: доля самоцитирований для любого из авторов статьи, а также по совокупности всех авторов статьи, не должна превышать 25 %; доля ссылок на статьи с участием одного автора, не являющегося автором (соавтором) статьи, не должна превышать 25 %.

6. Суммарная доля таблиц и иллюстраций в общем объеме представляемой статьи не должна превышать 40 %. Под иллюстрациями понимаются следующие объекты: диаграммы; графики; рисунки; эскизы; фотографии; карты и т.п.

7. Доля оригинального текста в статьях (оцениваемого через систему «Антиплагиат» на сайте www.antiplagiat.ru) должна быть не менее 80 %.

8. Указание на то, что работа финансируется по какому-либо гранту, в рамках Федеральной целевой программы, государственного заказа и пр. дается в виде постраничной сноски после заголовка (названия) работы.

9. В сведения об авторах работ помимо места работы и должности целесообразно включать ORCID автора и гиперссылку на страничку с его личными наукометрическими показателями на сайте www.elibrary.ru. По желанию можно привести также ссылки на странички с наукометрическими показателями на Scopus, в ResearchGate; на личную страничку, размещенную на сайте организации.

10. Основные технические требования к оформлению статей (материалов):

10.1. Текст должен быть расположен по ширине страницы формата А4 с учётом полей (все поля по 2,5 см), набран шрифтом Times New Roman, кегль 14, межстрочный интервал 1,0. В таблицах, подрисуночных надписях допускается уменьшенный шрифт – вплоть до 10 кегля. Альбомная ориентация страниц допускается только в порядке исключения для следующих случаев: широкоформатные таблицы с большим количеством колонок; иллюстрации большого размера, которые не умещаются на странице с книжной ориентацией.

Абзацные отступы одинаковы по всему тексту – 1,25 см. Кавычки («»), скобки ([ ], ( )), маркеры и другие знаки должны быть аналогичными на протяжении всего предоставляемого для публикации материала.

10.2. Все таблицы, рисунки, формулы должны иметь сквозную нумерацию в пределах текста статьи. Заголовки таблиц пишутся над ними и должны включать в себя номер таблицы и ее содержательное наименование.

10.3 Формулы должны иметь сквозную нумерацию в пределах текста статьи. Для формул желательно избегать «многоэтажных конструкций». Нумероваться могут только те формулы, на которые есть ссылки в статье. Размеры шрифтов для формул в MS Equation Editor: основные символы – 14 пт.; подстрочные и надстрочные индексы – 10 пт.; дополнительные индексы для подстрочных и надстрочных индексов – 8 пунктов. Константы в формулах записываются прямым шрифтом, переменные – курсивом (наклонным шрифтом); векторные и матричные величины – полужирным шрифтом. Матричные величины могут указываться в квадратных скобках, векторные величины (наборы скаляров) – в фигурных скобках.

Дополнительная информация для авторов статей

Основная тематика журнала и требования к научному уровню представляемых статей – http://hi-tech.asu.edu.ru/docs/ru-rules/ trebovaniya.pdf.

Адрес и телефоны редакционно-издательского дома – http://hi-tech.asu.edu.ru/docs/ru-rules/address.pdf.

Пример оформления текста статьи – http://hi-tech.asu.edu.ru/docs/ru-rules/primer.pdf.

Рекомендации по оформлению графических аннотаций – http://hi-tech.asu.edu.ru/docs/ru-rules/graf_ann.pdf, дополнительных материалов к статьям – http://hi-tech.asu.edu.ru/docs/ru-rules/dop_mat.pdf.

Подробные правила оформления таблиц – http://hi-tech.asu.edu.ru/docs/ru-rules/tables.pdf, оформления формул – http://hi-tech.asu.edu.ru/docs/ru-rules/formul.pdf, оформления иллюстраций – http://hi-tech.asu.edu.ru/docs/ru-rules/ilustr.pdf.

Пример оформления основного списка литературы к статье – http://hi-tech.asu.edu.ru/docs/ru-rules/spisok_liter.pdf, транслитерированного списка- – http://hi-tech.asu.edu.ru/docs/ru-rules/trasn_spisok_liter.pdf.

Дополнительные документы к статьям, представляемые авторами – http://hi-tech.asu.edu.ru/docs/ru-rules/dop_doc.pdf.

Переписка с авторами и порядок рецензирования статей – http://hi-tech.asu.edu.ru/docs/ru-rules/poryadok_recenz.pdf.

Условия и порядок оплаты статей – http://hi-tech.asu.edu.ru/docs/ru-rules/poryadok_oplaty.pdf.

Подписка на наши издания осуществляется
по Объединённому каталогу «Пресса России»

Журнал фундаментальных и прикладных исследований ***«Гуманитарные исследования»***
Подписной индекс – 11109

В журнале публикуются статьи по широкому спектру проблем гуманитарного знания. Ведущие направления публикаций отражены в следующих рубриках: «Языкознание», «Литературоведение».
Периодичность издания – 4 раза в год.
Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Научно-технический журнал ***«Геология, география и глобальная энергия»***
Подписной индекс – 11173

Редколлегия журнала принимает к рассмотрению статьи по проблемам геологии, нефтегазоностности различных регионов, охватывающие важнейшие и крайне полезные для науки и производства, а также для обучения студентов естественного направления.

Периодичность издания – 4 раза в год.
Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Научный журнал «***Каспийский регион: политика, экономика, культура»***
Подписной индекс – 11170

Профиль журнала – анализ проблем настоящего, прошлого и будущего Каспийского региона в их взаимосвязи с современным развитием мира.
Издание имеет многоплановый, междисциплинарный характер, знакомит читателя с исследованиями и дискуссиями во всех областях социальных и гуманитарных знаний по проблемам Каспийского региона.
Периодичность издания – 4 раза в год.
Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Научно-технический журнал ***«Прикаспийский журнал: управление и высокие технологии»***
Подписной индекс – 73313

На страницах журнала представлены результаты исследований и новейшие разработки в области технических наук.
Ведущие направления публикаций отражены в следующих рубриках: «Управление в социальных и экономических системах», «Системный анализ, управление и обработка информации», «Математическое моделирование, численные методы и комплексы программ», «Информационная безопасность и защита информации», «Информационно-измерительные и управляющие системы».
Периодичность издания – 4 раза в год.
Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Предлагаем всем желающим разместить в наших изданиях рекламу.
Адрес Астраханского государственного университета:
414056, г. Астрахань, ул. Татищева, 20а; тел. (8512) 24-64-95, 24-68-37
e-mail: asupress@yandex.ru

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2022
№ 1 (57)

Свидетельство о регистрации средства массовой информации
Федеральной службы по надзору в сфере массовых коммуникаций,
связи и охраны культурного наследия
ПИ № ФС77-31932 от 16 мая 2008 г.

Учредитель
Астраханский государственный университет
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20а

Адрес редакции:
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20

Адрес издателя:
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20а

Подписной индекс – 73313
По Объединенному каталогу «Пресса России»

Редактирование,
компьютерная правка, верстка *Н.Н. Сахно*

Дата выхода в свет 30.03.2022 г.

Цена свободная
Уч.-изд. 12,3. Усл. печ. л. 17,2.
Заказ № 4413. Тираж 500 экз. (первый завод – 22 экз.)

Астраханский государственный университет
414056, г. Астрахань, ул. Татищева, 20а
Тел. (8512) 24-64-95, тел./факс (8512) 24-68-37
E-mail: asupress@yandex.ru