АЗЕРБАЙДЖАНСКАЯ РЕСПУБЛИКА

На правах рукописи

РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ ОБЕСПЕЧЕНИЯ КИБЕРУСТОЙЧИВОСТИ КИБЕРФИЗИЧЕСКИХ СИСТЕМ

Специальность: 3338.01 – Системный анализ, управление и обработка информации (информационные технологии)

Отрасль науки: Технические науки

Соискатель: **Сухостат Людмила Валентиновна**

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени доктора наук

Баку – 2025

Диссертационная работа выполнена в Институте Информационных Технологий Министерства Науки и Образования Азербайджанской Республики.

Научный консультант: академик НАН Азербайджана,
доктор технических наук, профессор
Алгулиев Расим Магамед оглы

Официальные оппоненты: доктор технических наук, профессор
Мусаева Наиля Фуад гызы
доктор технических наук, профессор
Мустафаев Валех Азад оглы
доктор технических наук, доцент
Джабраилова Зарифа Гасым гызы
доктор технических наук, профессор
Исмайлов Балами Гасым оглы

Диссертационный совет ED 1.35 Высшей аттестационной комиссии при Президенте Азербайджанской Республики, действующий на базе Института Информационных Технологий Министерства Науки и Образования Азербайджанской Республики

Председатель диссертационного совета:
академик НАН Азербайджана,
доктор технических наук, профессор
_____ **Алгулиев Расим Магамед оглы**

Ученый секретарь диссертационного совета:
доктор философии по техническим наукам, доцент
_____ **Абдуллаева Фаргана Джаббар гызы**

Председатель научного семинара:
доктор технических наук
_____ **Муталлимов Муталлим Мирзаахмед оглы**

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность и степень разработанности темы. Современные информационные технологии являются важным средством развития государства и общества. Киберфизические системы (КФС) являются базовой технологией в реализации концепции Индустрия 4.0 и ключевым компонентом интеллектуальных объектов, которые широко используются в различных областях человеческой деятельности.

Ключевые технологические тенденции, лежащие в основе КФС, включают: «Интернет вещей» (Internet of Things), электрические сети, использующие смарт-технологии, производственные системы с интегрированным интеллектом, облачные вычисления и т.д. КФС являются основой для развития следующих сфер: смарт-производство, смарт-медицина, смарт-здания и инфраструктуры, интеллектуальные транспортные системы, мобильные системы, системы обороны и системы метеонаблюдения [4].

КФС представляют собой комбинацию двух технологий: информационные технологии (ИТ) киберуровня и операционные технологии (ОТ) физического уровня. Киберуровень состоит из серверов, систем баз данных, хостов и т. д. Физический уровень состоит из сенсоров, исполнительных механизмов, систем обратной связи и т. д., которые отвечают за управление производственными объектами.

КФС требует подключения многих устройств и систем для сбора и обмена огромными объемами данных на платформах ОТ и ИТ. Для этого существуют различные протоколы автоматизации, которые необходимо соединить для достижения поставленных задач. Стандарт ISO 22301:2019 имеет решающее значение для повышения устойчивости организаций к различным непредвиденным сбоям, обеспечивая непрерывность операций и услуг.

Актуальность к обеспечению кибербезопасности КФС резко возросла в последние годы после серии кибератак.

Кибератаки на КФС представляют значительный риск для здоровья и безопасности человеческой жизни, угрожают серьезным ущербом окружающей среде. Интерес к обеспечению информационной безопасности КФС резко возрос после серии атак посредством компьютерного вируса Stuxnet в 2010 году. Из-за постоянно меняющихся кибер-угроз необходима разработка методологии более высокого уровня, которая требует обнаружения внедренных и скрывающихся кибератак, обеспечивая более киберустойчивые и защищенные киберинфраструктуры.

12 марта 2024 года Европейский парламент одобрил предложенный Закон о киберустойчивости (Cyber Resilience Act, CRA). Он будет применяться ко всем продуктам с цифровыми элементами, представленным на европейском рынке, использование которых включает прямое или косвенное логическое или физическое подключение данных к устройству или сети.

Нарушение функционирования КФС и вывод из строя ее компонентов может привести к серьезному ущербу, что ставит вопрос о необходимости рассмотрения кибербезопасности КФС как одного из приоритетных.

Переход к Индустрии 4.0 поставил перед Азербайджанской Республикой, как и перед всем миром, новые задачи и вызовы. В Азербайджане обеспечены все возможности для развития Четвертой промышленной революции. В стране постоянно ведутся разработки передовых технологий. 19 марта 2025 года утверждена «Стратегия искусственного интеллекта на 2025-2028 годы».

На сегодняшний день в Республике все технологии Индустрии 4.0 применяются в различных критических инфраструктурах, таких как нефтегазовые платформы, транспортные системы и т.д.

Активное сотрудничество с Всемирным экономическим форумом привело к созданию в Баку Центра анализа и координации четвертой промышленной революции согласно

указу Президента Азербайджанской Республики от 6 января 2021 года. Деятельность центра призвана обеспечить развитие искусственного интеллекта и машинного обучения, а также служить гражданам, развитию цифровой экономики и дальнейшему совершенствованию различных областей исследований в стране.

17 апреля 2021 года был подписан указ «О некоторых мерах по обеспечению безопасности критической информационной инфраструктуры».

Согласно закону Азербайджанской Республики о внесении изменений в Закон Азербайджанской Республики «Об информации, информатизации и защите информации» от 27 мая 2022 года добавляются понятия «критическая информационная инфраструктура», ее объект и субъект, а также правовая основа правил обеспечения ее безопасности. 28 августа 2023 утверждена «Стратегия информационной и кибербезопасности Азербайджанской Республики на 2023-2027 годы». 29 ноября 2024 года кабинет министров Азербайджана утвердил «Перечень объектов критической информационной инфраструктуры».

Указом Президента Азербайджанской Республики от 16 января 2025 года утверждена «Концепция цифрового развития в Азербайджанской Республике».

Выявление современных кибератак имеет решающее значение для устойчивой работы КФС. Традиционные механизмы защиты становятся все более неэффективными из-за методов, используемых злоумышленниками. Риски кибератак возникают из-за интеграции кибер- и физического доменов, другими словами, интеграции доменов ИТ и ОТ. В этом контексте разработка адаптивных, более эффективных методов обнаружения кибератак является насущной потребностью для защиты ИТ и ОТ инфраструктур от таких угроз.

Критические ситуации идеальны для киберпреступников, поскольку они могут воспользоваться самым слабым звеном – человеческим фактором. Киберпреступники используют

человеческий фактор для получения несанкционированного доступа, кражи учетных данных и заражения систем вредоносными программами.

Важно заранее разработать алгоритмы обнаружения и меры противодействия всем известным кибератакам, чтобы уменьшить их влияние и минимизировать ущерб системе. Однако исследования показывают, что требуется более глубокий анализ кибербезопасности КФС, особенно применимости подходов на основе искусственного интеллекта.

В первую пятерку исследовательских университетов, занимающихся этой проблемой, входят Калифорнийский университет в Беркли (США), Пекинский университет науки и технологий (Китай), Королевский технологический институт KTH (Швеция), Миланский политехнический университет (Италия) и Гамбургский технологический университет (Германия).

По мере продвижения к Индустрии 4.0 КФС становится все более сложной, распределенной и управляемой данными, что постепенно превращает обнаружение вторжений и аномалий в проблему больших данных. Сложной задачей становится необходимость анализа большого объема данных, разнородных по своей природе, с минимальными вычислительными затратами.

Использование сенсоров, исполнительных механизмов и других устройств для сбора данных из КФС предоставляет полезную информацию для их последующего анализа. Однако сегодня объем хранимых данных становится слишком большим для обработки традиционными алгоритмами. Исследователям приходится прибегать к различным подходам, таким как разбиение данных и использование априорных знаний. Таким образом, работа с большими данными вызывает необходимость формализации новых методов, применяемых исследователями, и создания новых алгоритмов, использующих возможности методов машинного обучения для работы с данными больших объемов и размерностей. Тем не менее, существующие

алгоритмы не всегда эффективны из-за высокой вычислительной сложности.

Методы машинного обучения широко применяются для анализа данных КФС на предмет наличия кибератак. Теоретически алгоритмы машинного обучения могут получить высокую производительность, т.е. минимизировать уровень ложных тревог и максимизировать точность обнаружения [10]. Однако обычно требуется бесконечное число обучающих образцов. На практике это условие невозможно в силу ограничения вычислительной мощности и требования ответа в режиме реального времени [10]. В связи с этим злоумышленники часто ищут и нацеливаются на самое слабое звено – наиболее уязвимый функциональный компонент в КФС. Самые уязвимые устройства могут стать хорошей отправной точкой для эффективного исследования уязвимостей системы безопасности. Существующие методы не могут всесторонне анализировать распространенные многоэтапные кибератаки в среде КФС [30].

Решение исследуемых проблем является важной задачей, в связи с этим были разработаны и применяются на практике различные методы и алгоритмы. Большинство из них рассматривают либо кибербезопасность только в среде ИТ, либо защищенность в среде ОТ. Обеспечение защиты от киберфизических атак является серьезной проблемой, требующей методов оценки киберустойчивости, способных исследовать тесные взаимодействия и взаимозависимости между кибер- и физическими компонентами в КФС. Однако существующие методы не учитывают эту специфику.

В свете вышеизложенного можно сделать вывод, что на данный момент не существует единого подхода к проектированию защищенных и кибербезопасных КФС. Данное диссертационное исследование направлено на разработку методов и алгоритмов для обеспечения кибербезопасности компонентов ИТ и защищенности компонентов ОТ.

Объект и предмет исследования. Объектом исследования является кибербезопасность ИТ и защищенность ОТ киберфизических систем, **предметом исследования** являются модели и методы управления процессами обеспечения киберустойчивости киберфизических систем.

Цель и задачи исследования. Целью диссертационной работы является разработка методов и алгоритмов обеспечения киберустойчивости и информационной безопасности киберфизических систем на основе интеллектуального анализа больших данных о кибератаках и отказах систем.

Для достижения поставленной цели в работе решаются следующие **задачи:**

- Анализ и исследование существующих методов и алгоритмов обеспечения киберустойчивости КФС;
- Разработка концептуальной модели обеспечения киберустойчивости КФС;
- Разработка методов и алгоритмов интеллектуальной обработки больших данных для обеспечения киберустойчивости КФС;
- Разработка методов и алгоритмов обеспечения кибербезопасности информационных технологий КФС;
- Разработка методов обеспечения защищенности операционных технологий КФС;
- Разработка метода определения критичности уязвимостей функциональных компонентов для обеспечения киберустойчивости КФС;
- Разработка метода оценки рисков КФС.

Методы исследования базируются на применении теории информационной безопасности, машинного обучения, теории вероятностей, теории нечетких чисел, теории распознавания образов, анализа рисков и принятия решений.

Основные положения выносимые на защиту

- Концептуальная модель обеспечения киберустойчивости КФС;

- Алгоритм обнаружения выбросов в больших данных на основе метода к-средних для снижения риска блокировки КФС;
- Алгоритм анализа больших данных КФС на основе взвешенной кластеризации;
- Метод параллельной обработки больших данных для снижения риска выхода КФС из строя;
- Метод кластерного анализа больших данных КФС на основе консенсусного ансамбля;
- Метод классификации кибератак на КФС с применением экстремального машинного обучения;
- Метод обнаружения DoS атак на КФС с применением ансамбля классификаторов;
- Алгоритм обнаружения вредоносного программного обеспечения в КФС на основе изображений;
- Метод обнаружения аномалий в операционных технологиях КФС на основе иерархических скрытых Марковских моделей для системы водоочистки;
- Метод обнаружения аномалий в акустических сигналах КФС с применением трансферного обучения;
- Метод обнаружения атак на устройства КФС с применением глубокой гибридной модели для системы водоочистки;
- Метод классификации отказов КФС на основе изображений;
- Метод определения критичности уязвимостей функциональных компонентов КФС на основе Байесовского графа атак для транспортировки природного газа по трубопроводу;
- Метод оценки рисков КФС с применением нечеткого интеграла Сугено для генерации энергии ветра.

Научная новизна исследования состоит в следующем:

- Разработана трехуровневая концептуальная модель для решения ряда задач по обеспечению киберустойчивости КФС;
- Разработан алгоритм обнаружения выбросов в больших данных на основе метода к-средних для снижения риска блокировки КФС;
- Разработан алгоритм анализа больших данных КФС на основе взвешенной кластеризации;
- Разработан метод параллельной обработки больших данных для снижения риска выхода КФС из строя;
- Разработан метод кластерного анализа больших данных КФС на основе консенсусного ансамбля;
- Разработан метод классификации кибератак на КФС с применением экстремального машинного обучения;
- Разработан метод обнаружения DoS атак на КФС с применением ансамбля классификаторов;
- Разработан алгоритм обнаружения вредоносного программного обеспечения в КФС на основе изображений;
- Разработан метод обнаружения аномалий в операционных технологиях КФС на основе иерархических скрытых Марковских моделей для системы водоочистки;
- Разработан метод обнаружения аномалий в акустических сигналах КФС с применением трансферного обучения;
- Разработан метод обнаружения атак на устройства КФС с применением глубокой гибридной модели для системы водоочистки;
- Разработан метод классификации отказов КФС на основе изображений;
- Разработан метод определения критичности уязвимостей функциональных компонентов КФС на основе

Байесовского графа атак для транспортировки природного газа по трубопроводу;

- Разработан метод оценки рисков КФС с применением нечеткого интеграла Сугено для генерации энергии ветра.

Теоретическая и практическая значимость исследования. Научная и практическая ценность рассматриваемых в диссертации задач заключается в том, что разработанные методы и алгоритмы могут быть использованы для обеспечения киберустойчивости киберфизических систем с целью повышения защищенности ОТ и кибербезопасности ИТ. Разработанные методы и алгоритмы опубликованы в научных журналах и представлены в виде докладов на международных научных конференциях и являются общедоступными для применения. Научная обоснованность и достоверность полученных результатов подтверждаются применением современных методов исследований, а также сравнением с результатами ранее предложенных методов и алгоритмов в исследуемой области с целью подтверждения их эффективности.

Полученные в диссертационной работе результаты имеют практическое значение и могут быть использованы в следующих областях:

- разработка предложений по стратегиям и программам обеспечения кибеустойчивости КФС;
- в системах обнаружения кибератак с целью обеспечения надежности, кибербезопасности, киберустойчивости и эффективности КФС;
- применение экспертами при анализе киберустойчивости КФС с целью повышения защищенности ОТ и кибербезопасности ИТ.

Апробация и внедрение. Основные научно-теоретические и практические результаты диссертации были представлены и обсуждались на: семинарах Института Информационных Технологий Министерства Науки и Образования АР, а также на следующих республиканских и международных научных конференциях: III Республиканской научно-практической конфе-

ренции «Riyaziyyatın tətbiqi məsələləri və yeni informasiya texnologiyaları» (г. Сумгаит, 15-16 декабря 2016 г.); I Республиканской конференции «Proqram mühəndisliyinin aktual elmi-praktiki problemləri» (г. Баку, 17 мая 2017 г.); I Республиканской научно-практической конференции «Big data: imkanları, multidissiplinar problemləri və perspektivləri» (г. Баку, 25 февраля 2016 г.); XIII Международной научно-технической конференции «Распознавание-2017» (г. Курск, Россия, 16-19 мая 2017 г.); IV Республиканской конференции «İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri» (г. Баку, 14 декабря 2018 г.); X Международной конеренции «Application of Information and Communication Technologies» (AICT-2016) (г. Баку, 12-14 октября 2016 г.); III Республиканской научно-практической конференции «İnformasiya təhlükəsizliyinin aktual problemləri» (г. Баку, 8 декабря 2017 г.); XV Международной научно-технической конференции «Распознавание-2019» (г. Курск, Россия, 14-17 мая 2019 г.); XIII Международной молодежной научной конференции «Technical Sciences. Industrial Management» (г. Боровец, Болгария, 11-14 марта 2020 г.); «Global Cyber Security Forum 2019» (г. Харьков, Украина, 14-16 ноября 2019); XIV Международной научно-технической конференции «Распознавание-2018» (г. Курск, Россия, 25-28 сентября 2018 г.); XIII Международной конеренции «Application of Information and Communication Technologies» (AICT-2019) (г. Баку, 23-25 октября 2019 г.); V Республиканской конференции «İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri» (г. Баку, 29 ноября 2019 г.); Международной конференции «Information Security: Problems and Prospects» (г. Баку, 29 октября 2021 г.); Международном научном семинаре NATO ARW «Cybersecurity of Industrial Control Systems (ICS)» (г. Баку, 27-29 октября 2021 г.); XI «Национальном Суперкомпьютерном Форуме (НСКФ-21)» (г. Переславль-Залесский, Россия, 30 ноября – 3 декабря 2021 г.); XVI Международной конференции «Application of Information and Communication Technologies» (AICT-2022) (г. Вашингтон, США, 12-14 октября 2022 г.); XVII

Международной научно-технической конференции «Распознавание-2023» (г. Курск, Россия, 12-15 сентября 2023 г.).

Соискатель принимал участие в конкурсах грантов, проводимым Фондом Развития Науки при Президенте Азербайджанской Республики (№ EİF-11-1(3)-82/08/1, EİF-RİTN-MQM-2/İKT-2-2013-7(13)-29/18/1, EİF-KETPL-2-2015-1(25)-56/05/1, AEF-MCG-2023-1(43)-13/04/1-M-04), Научным фондом Государственной нефтяной компании Азербайджанской Республики (ГНКАР) и Национальной Академией Наук Азербайджана.

Название организации, в которой выполнена диссертационная работа. Диссертация была выполнена в Институте Информационных Технологий Министерства Науки и Образования Азербайджанской Республики.

По теме диссертационной работы опубликованы 33 научных работы, из них 17 статей опубликованы в зарубежных журналах, в том числе 2 статьи в рецензируемых журналах, рекомендованных Высшей аттестационной комиссией при Президенте Азербайджанской Республики, 4 статьи входят в международную базу данных Scopus и 11 работ входят в базу данных Web of Science. 16 статей опубликованы в материалах международных и республиканских конференций.

Объем и структура работы. Диссертационная работа состоит из введения, пяти глав, заключения и списка литературы (409 наименований, в том числе 11 на азербайджанском языке, 13 на русском и 385 на английском языке). Основное содержание работы изложено на 318 страницах, включая 54 рисунка и 80 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулированы цель и задачи, приведены научная новизна, теоретическая и практическая ценность работы и основные положения, выносимые на защиту, структура и содержание работы.

Первая глава посвящена анализу и исследованию существующих методов и алгоритмов обеспечения киберустойчивости и информационной безопасности КФС. Создание КФС систем поставило перед людьми новые задачи. Обеспечение информационной безопасности КФС является одной из наиболее сложных проблем в широком спектре средств защиты от кибератак. В данной главе описывается принцип работы КФС. Рассматриваются основные виды кибератак и угроз КФС. Проводится анализ существующих исследовательских работ по интеллектуальной обработке больших данных КФС, обеспечению кибербезопасности ИТ, обнаружению аномалий в ОТ и оценке киберустойчивости КФС. Глава состоит из шести параграфов.

В первом параграфе проводится анализ особенностей становления и развития КФС. Описана общая структура КФС, представлены основные стандарты, которые могут служить полезным руководством по оценке киберустойчивости КФС. Для анализа последних исследований в области обеспечения кибербезопасности КФС были выделены четыре категории исследований: оценка последствий кибератак, моделирование атак на КФС, обнаружение атак на КФС и разработка архитектуры безопасности.

Во втором параграфе проводится анализ методов и алгоритмов интеллектуальной обработки больших данных КФС. Использование сенсоров, исполнительных механизмов и других устройств для сбора данных КФС предоставляет полезную информацию для их последующего анализа. Работа с большими данными вызывает необходимость формализации новых методов, применяемых исследователями, и создания новых алгорит-

мов, использующих возможности методов машинного обучения. Исследователям приходится прибегать к различным подходам, таким как разбиение данных и использование априорных знаний на предмет снижения риска, реализации DoS атак и блокировки функционирования ИТ КФС. Тем не менее, существующие алгоритмы не всегда эффективны из-за высокой вычислительной сложности.

Третий параграф посвящен анализу и исследованию методов обеспечения кибербезопасности ИТ и защищенности ОТ КФС. Описаны некоторые современные решения в области анализа рисков, обнаружения вторжений, киберустойчивости и реагирования на инциденты в КФС. Традиционные механизмы защиты становятся все менее эффективными из-за их использования злоумышленниками. В связи с этим разработка более эффективных методов обнаружения кибератак является важнейшей задачей защиты ИТ и ОТ КФС от киберугроз. По результатам анализа существующих исследований в области безопасности предлагается «дерево атак» и угроз на основе функциональной модели КФС. Ветви «дерева» включают следующие виды атак: а) атаки на сенсорные устройства; б) атаки на исполнительные механизмы; в) атаки на вычислительные компоненты; г) атаки на коммуникации; д) атаки на обратную связь. Обнаружение аномалий в промышленном сценарии имеет важное значение, поскольку необнаруженные сбои могут привести к критическим повреждениям. Раннее обнаружение аномалий может повысить надежность подверженного сбоям промышленного оборудования и снизить затраты на эксплуатацию и техническое обслуживание.

В четвертом параграфе проанализированы и исследованы методы и алгоритмы оценки киберустойчивости КФС. В настоящее время предложено множество подходов к анализу киберустойчивости КФС, однако мало внимания уделяется корреляции между устройствами и уязвимостями. Существующие методы не могут всесторонне анализировать распространенные многоэтапные кибератаки в среде КФС. Большинство современных работ сосредоточено в основном на теоретических исследованиях и

содержит лишь краткую оценку киберустойчивости КФС на основе методов машинного обучения. Граф уязвимостей – многообещающий метод, который перечисляет все возможные пути кибератак с помощью серии эксплойтов. Следовательно, исследователи и практики проявляют интерес к анализу киберустойчивости КФС на основе графов. Целью киберустойчивости является всеобъемлющая защита всей КФС с охватом всех доступных киберресурсов.

В пятом параграфе дается постановка научной задачи исследования. Приводятся основные понятия и определения, используемые в диссертационной работе. При этом основное внимание уделяется описанию понятия киберустойчивость КФС. Рассматриваются основные виды кибератак и киберугроз КФС.

Шестой параграф посвящен разработке концептуальной модели для решения ряда задач по обеспечению киберустойчивости КФС. Предлагаемая концептуальная модель – это иерархическая архитектура, которая состоит из трех основных уровней: уровень краевых вычислений, уровень туманных вычислений и уровень облачных вычислений (Рис. 1). Уровень краевых вычислений включает в себя такие интеллектуальные устройства как сенсоры и исполнительные механизмы, которые используются для сбора данных и их предварительной обработки в режиме реального времени. Этот уровень временно хранит данные и может генерировать первоначальные решения об обнаружении аномального состояния. На уровне краевых вычислений также выполняется аутентификация кибербезопасности на основе протоколов. Несмотря на то, что данный уровень выполняет предварительную обработку данных сенсоров и исполнительных механизмов в режиме реального времени, необходима их дополнительная обработка для уточнения полученных результатов и последующего принятия решений. С этой целью полученные данные передаются на уровень туманных вычислений.

Уровень туманных вычислений состоит из большого количества распределенных узлов, содержащих шлюзы, сервисы туманных

вычислений, маршрутизаторы, точки доступа, базовые станции и коммутаторы. Туманные вычисления предоставляют децентрализованную платформу, направленную на обнаружение аномальной работы КФС, отказов системы, вызванных кибератаками и обнаружение вредоносного ПО, и расширяют облачные сервисы до обработки и анализа границ сети.

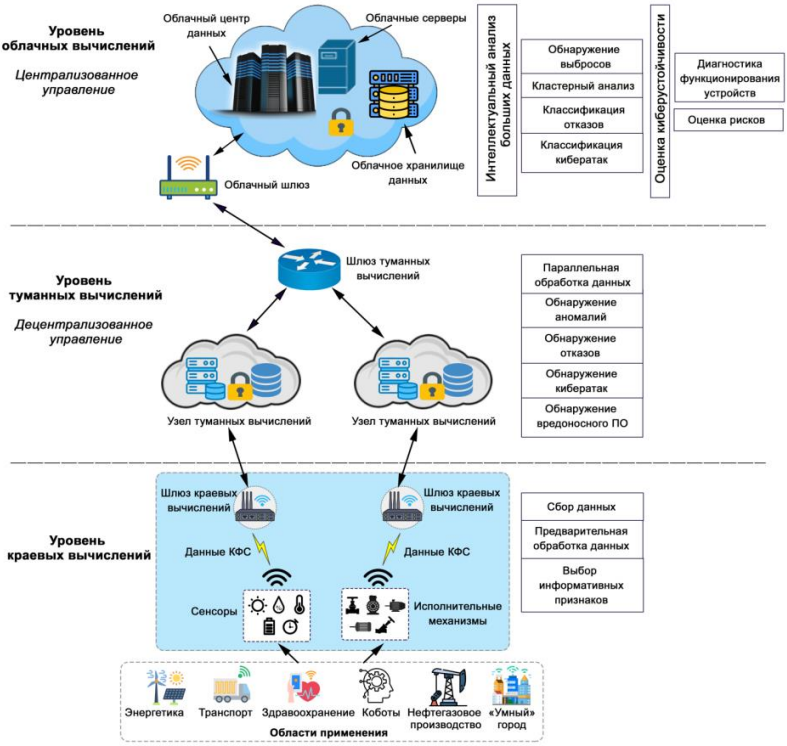


Рис. 1. Концептуальная модель обеспечения киберустойчивости КФС

Уровень туманных вычислений расположен на границе сети, где узлы туманных вычислений размещены на конечных устройствах и облачных слоях. Это позволяет снизить вычислительную сложность и потребление энергии, обеспечить использование меньшего количества ресурсов и высокую эффективную мощ-

ность. При этом вычисления могут быть статичными, а также динамичными (мобильными).

Уровень облачных вычислений состоит из нескольких серверов, хранилища данных, облачного центра данных и шлюзов. Данный уровень производит высокопроизводительные вычисления и предоставляет услуги для таких областей как энергетика, транспорт, здравоохранение, коботы, нефтегазовое производство, «умный» город и т.д. Уровень облачных вычислений осуществляет централизованное управление, обеспечивает постоянное хранение больших объемов данных и осуществляет их интеллектуальный анализ, в том числе оценку киберустойчивости КФС. Интеллектуальный анализ больших данных предоставляет возможность потоковой и пакетной обработки. Это приводит к минимуму действия человека в производственном процессе, а экспертная оценка данных (параллельная обработка, обнаружение выбросов, кластерный анализ, классификация отказов и кибератак) может проводиться в режиме реального времени небольшой группой экспертов или с применением технологий машинного обучения. Устройства облачных вычислений подключены к узлам туманных вычислений с помощью проводного соединения и беспроводных сред (Bluetooth, 5G, ZigBee, WiFi и беспроводной локальной сети). Уровень облачных вычислений производит обучение больших и сложных моделей на основе глубоких нейронных сетей, которые могут передаваться на узлы краевых и туманных вычислений по запросу. Это позволяет снизить потребление вычислительных ресурсов на уровнях туманных и краевых вычислений.

Предлагаемая архитектура может предоставить необходимое руководство исследователям, промышленникам и специалистам при анализе каждого слоя архитектуры КФС с точки зрения моделирования системы, управления, мониторинга, сбора данных и их анализа для достижения требуемой надежности, кибербезопасности, киберустойчивости и эффективности.

Таким образом, были исследованы особенности становления и развития КФС. Проведен анализ методов и алгоритмов интел-

лектуальной обработки больших данных КФС, обеспечения кибербезопасности ИТ и защищенности ОТ КФС, а также оценки киберустойчиво-сти КФС. Анализ существующих методов управления процессами обеспечения киберустойчивости КФС показывает, что актуальными являются следующие проблемы: разработка методов и алгоритмов интеллектуальной обработки больших данных для обеспечения киберустойчивости КФС; разработка методов и алгоритмов обеспечения кибербезопасности ИТ КФС и защищенности ОТ КФС; разработка метода определения критичности уязвимостей функциональных компонентов КФС и метода оценки рисков КФС.

Вторая глава посвящена разработке методов и алгоритмов интеллектуальной обработки больших данных для обеспечения киберустойчивости КФС. При анализе данных качество информации имеет первостепенное значение. Эта задача усложняется ростом больших объемов собираемой информации. Работа с большими данными требует больших вычислительных ресурсов. В связи с этим исследователи уделяют особое внимание разработке эффективных методов и алгоритмов интеллектуального анализа больших данных. Высокая степень важности решаемых задач привела к тому, что в этой области появилась целая плеяда различных методов. Методы отличаются друг от друга простотой реализации, пригодностью для обработки данных и лежащими в их основе базовыми принципами. Глава состоит из четырех параграфов.

В первом параграфе представлен алгоритм обнаружения выбросов в больших данных на основе метода k-средних для снижения риска блокировки КФС.

Аналитика больших данных требует больших вычислительных ресурсов и ресурсов памяти. Поскольку эти ресурсы не всегда доступны и требуют больших затрат, предложение новых алгоритмов анализа больших данных считается одним из наиболее экономически эффективных подходов. Алгоритм k-средних имеет быструю скорость сходимости и не требует больших вычислительных ресурсов. Поэтому он популярен при кластериза-

ции небольших наборов данных, но требует больших вычислительных затрат при увеличении размеров наборов данных.

Пусть $x_i \in R^n$ $(i = \overline{1,n})$ – точка из набора данных, где n – общее количество точек во входном наборе данных, $x_i \in c_p \in R^k$ $(p = \overline{1,k})$ – номер кластера, где k – количество кластеров, S_W – компактность кластеров, S_{BW} – разделимость кластеров друг от друга и S_B является мерой удаленности центра каждого кластера (O_p) от центра всех точек (O) во входном наборе данных

$$S_W = \sum_{p=1}^{k} \sum_{i=1}^{n} (x_i - O_p)(x_i - O_p)^T, \quad (1)$$

$$S_{BW} = \sum_{p=1}^{k-1} \sum_{q=p+1}^{k} (O_p - O_q)(O_p - O_q)^T, \quad (2)$$

$$S_B = \sum_{p=1}^{k} (O_p - O)(O - O_p)^T, \quad (3)$$

$$O = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad O_p = \frac{1}{n_p} \sum_{x_i \in C_p} x_i, \quad n_p = |C_p|, \quad p = 1,2,\dots,k. \quad (4)$$

Предложены три алгоритма для обнаружения выбросов в больших данных КФС с применением кластеризации. Алгоритмы основаны на кластеризации и учитывают компактность и разделимость кластеров. Для первого предложенного алгоритма задача состоит в том, чтобы максимизировать следующую функцию:

$$F_1(x) = \frac{S_B + S_{BW}}{S_W} \longrightarrow \max. \quad (5)$$

Для второго алгоритма задача состоит в том, чтобы максимизировать следующую функцию:

$$F_2(x) = \frac{S_B * S_{BW}}{S_W} \longrightarrow \max. \quad (6)$$

В третьем алгоритме задача состоит в максимизации целевой функции по параметру регуляризации (α), который определяется экспериментально:

$$F_3(x) = \frac{1}{S_W} (\alpha S_B + (1 - \alpha) S_{BW}) \longrightarrow \max. \quad (7)$$

Алгоритмы минимизируют компактность кластеров и максимально отделяют кластеры друг от друга по расстояниям между их центроидами и удаляют центры кластеров от выбранного общего центра точек в наборе данных.

Численные эксперименты проводились на малых, средних и больших реальных наборах данных и продемонстрировали эффективность предложенных алгоритмов. Рассматривались три набора, включая наборы данных Phishing и Spam из репозитория UCI, а также набор данных NSL-KDD. Набор данных сигнатур атак NSL-KDD построен на основе базы данных KDD-99. Он содержит обучающую (125973 образцов) и тестовую (22544 образцов) выборки. Метки назначаются каждому экземпляру либо как «аномалия», либо как «нормальное» состояние. Все образцы (148517) рассматривались в данном исследовании. Набор данных Spam содержит электронные письма со спамом и без него. Он включает в себя признаки, которые указывают как часто встречалось конкретное слово или символ в электронном письме, и измеряют длину последовательностей заглавных букв. Набор данных содержит два класса: спам («аномалия») или «нормальное» состояние. Phishing содержит 11055 фишинговых веб-сайтов. Он включает в себя 30 признаков (IP-адрес, длины URL-адреса и аномального URL-адреса, перенаправления вебсайта и т. д.). Наборы данных были разделены на два класса. В ходе предварительной обработки значения в наборах данных были стандартизированы.

Было рассмотрено влияние параметра регуляризации α на эффективность третьего предложенного алгоритма при различных наборах данных. Для оценки эффективности предложенных алгоритмов были рассмотрены метрики кластеризации на основе расстояния между парами точек данных, являются «чистота» (purity), метрика Миркина (Mirkin), коэффициент разделения (partition coefficient, PC), вариация информации (variation of information, VI), F-мера и V-мера. Полученные результаты для первого, второго, третьего предложенных алгоритмов и алгоритма k-средних более наглядно проиллюстрировано на рис. 2. Сравнение с алгоритмом k-средних доказало превосходство предложенного подхода. На основании экспериментальных результатов можно сделать вывод, что три предложенных алгоритма превосходят алгоритм k-средних по метрикам Purity, F-

мера и PC для набора данных Phishing. Для набора данных NSL-KDD наилучшие результаты были получены при применении первого и второго предложенных алгоритмов согласно приведенным выше трем метрикам. По всем метрикам второй алгоритм показал лучший результат для набора данных Phishing. При этом метрики Purity, Mirkin, F-мера и PC показали хорошие результаты для набора данных NSL-KDD при применении второго алгоритма. Третий предложенный алгоритм показал наилучшие результаты для наборов данных NSL-KDD и Spam, а для набора данных Phishing наблюдались самые низкие значения. Можно сделать вывод, что третий алгоритм хорошо работает на наборах данных малого и большого размера, тогда как второй алгоритм показал наилучшие результаты на наборах данных среднего размера.

Численные эксперименты на наборах данных различной размерности продемонстрировали эффективность предложенных алгоритмов.

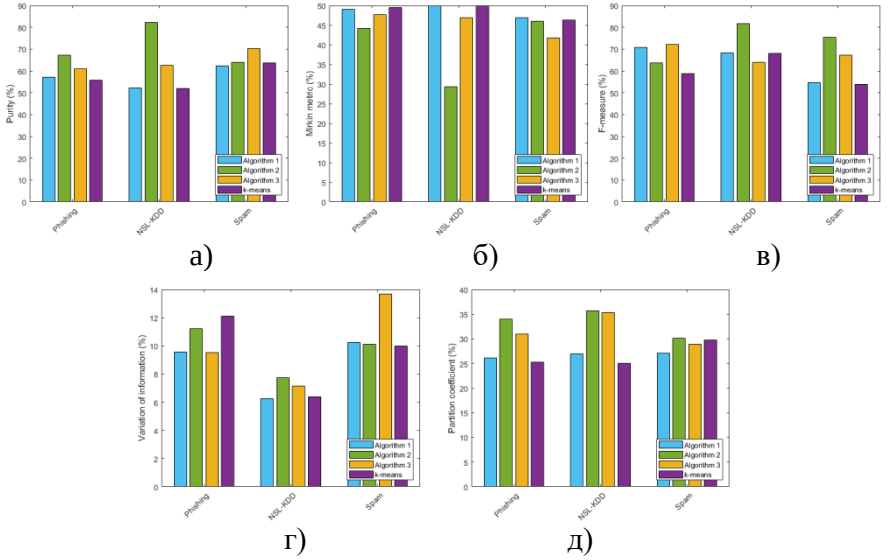


Рис. 2. Сравнение предложенных алгоритмов с алгоритмом k-средних на основе метрик

Сравнение с алгоритмом k-средних доказало превосходство предложенных алгоритмов.

Во втором параграфе представлен алгоритм анализа больших данных КФС на основе взвешенной кластеризации. Данное исследование направлено на разработку кластерного подхода для обнаружения аномалий в больших данных Веса, полученные путем суммирования весов каждой точки из набора данных, были назначены кластерам. Взвешивание применяется для улучшения решения кластеризации.

Пусть $X = (x_1, x_2, \dots, x_n)$ – точки в наборе данных, где n - общее количество точек в наборе данных, $x_i = (x_{i1}, x_{i2}, \dots, x_{im}) \in R^m$ – точка в наборе данных, где m - размерность точек данных, $C = (C_1, C_2, \dots, C_k)$ – кластеры, где C_p ($p = \overline{1,k}$) – p-ый кластер и k – количество кластеров. Задача состоит в том, чтобы минимизировать функцию для обнаружения аномалий в наборе данных следующим образом:

$$f(x) = \sum_{p=1}^{k} \sum_{x_i \in C_p} \left|C_p\right|_W * \left\|x_i - O_p\right\|^2 \longrightarrow \min, \qquad (8)$$

где $\left|C_p\right|_W$ – вес -го кластера.

В этом случае вес кластера определяется как сумма весов всех точек кластера:

$$\left|C_p\right|_W = \sum_{x_i \in C_p} w(x_i),\ p = 1,2, \dots, k, \qquad (9)$$

где веса точек рассчитываются на основе их расстояния от центра всех точек в наборе данных

$$w(x_i) = \|x_i - O\|,\ O = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad (10)$$

а центр -го кластера (O_p) определяется как

$$O_p = \frac{1}{n_p}\sum_{x_i \in C_p} x_i\ ,\ n_p = \left|C_p\right|,\ p = 1,2, \dots, k. \qquad (11)$$

В алгоритме каждый кластер представлен своим центром, и цель состоит в том, чтобы найти решение, которое минимизирует расстояние между каждой точкой и центром кластера, которому она назначена.

Результаты экспериментов показали эффективность предложенного подхода, как для кластеризации, так и при обнаруже-

нии аномалий одновременно. Сравнение с алгоритмом k-средних показало, что предложенный алгоритм точнее обнаруживает аномалии в КФС (Таблица 1). Для оценки производительности предложенного подхода использовалось относительное улучшение:

$$\frac{\text{предложенный\_метод} - \text{метод\_k-средних}}{\text{метод\_k-средних}} \times 100\%.$$

Таблица 1

Сравнение производительности алгоритма к-средних и предложенного алгоритма

Метрики / Набор данных	Purity (%)	Mirkin (%)	F-мера (%)	VI (%)	PC (%)
NSL-KDD	3.78 (+)	0.44 (+)	0.66 (+)	13.03 (-)	10.55 (+)
Phishing	5.58 (+)	2.18 (+)	1.21 (-)	17.72 (+)	8.88 (+)

На основании экспериментальных результатов можно сделать вывод, что предложенный подход превосходит алгоритм k-средних по четырем метрикам (Purity, Mirkin, F-мера и PC) для набора данных NSL-KDD. Purity, Mirkin, PC и VI показали хорошие результаты для набора данных Phishing. Значения показателей F-мера были достаточно близкими для обоих подходов к наборам данных NSL-KDD и Phishing. Предложенный подход становится более эффективным с увеличением размера анализируемого набора данных.

Результаты экспериментов показали эффективность предложенного подхода, как для кластеризации, так и при обнаружении аномалий одновременно. Экспериментальные результаты показали, что «взвешивание» улучшает решение кластеризации и предложенный алгоритм точнее обнаруживает аномалии в КФС по сравнению с алгоритмом k-средних.

В третьем параграфе представлен метод параллельной обработки больших данных для снижения риска выхода КФС из строя. Анализ наборов данных большой размерности требует наличия больших вычислительных мощностей, что не всегда осуществимо. Данное исследование направлено на решение

проблемы ускорения процесса кластеризации. Это было достигнуто за счет связывания параллелизации и разделения набора данных на небольшие пакеты (батчи) с использованием кластеризации на основе k-средних.

Пусть $X = \{x_1, x_2, \dots, x_n\}$ – конечное число точек, заданных в -мерном пространстве, q – размер батча, максимальное значение $q(q < n)$ которого определяется параметрами ПК, а также обрабатывается за разумное время, $C = \{C_1, C_2, \dots, C_k\}$ – множество кластеров, где $C_p^q (p = 1,2, \dots, k)$ — p-ый кластер q-го батча, а k – количество кластеров, O_p^q — центроид p-го кластера в q-ом батче.

Целевая функция имеет следующий вид:

$$f(x) = \sum_{p=1}^{k} \sum_{x_i \in C_p} \left\| x_i - O_p^q \right\|^2 \longrightarrow \min, \qquad (12)$$

$$O_p^q = \frac{\sum_{x_i \in C_p^q} x_i}{\left| C_p^q \right|}, \; p = 1,2, \dots, k, \qquad (13)$$

где $\|.\|$ - Евклидова норма в $\mathfrak{R}^m$, $\left| C_p^q \right|$ – число точек данных в кластере C_p^q.

Результирующий центроид, полученный после применения алгоритма k-средних к набору центроидов всех батчей, обозначается как $O_p^* \; (p = 1,2, \dots, k)$. В этом случае набор данных разбивается на несколько равных батчей. Результирующие кластеры создаются параллельно без полной загрузки памяти, что значительно ускоряет кластеризацию. Метод работает параллельно и итеративно. Актуальность исследования заключается в том, что использование батчей малой размерности снижает вычислительные затраты и увеличивает скорость сходимости алгоритма кластеризации.

Эксперименты проводились на двух наборах данных: Phone Accelerometer и Individual Household Electric Power Consumption. Набор данных Phone Accelerometer содержит 1048575 образцов и 4 признака. А для второго набора данных количество образцов составило 2075259 и 9 признаков. В экспериментах эти наборы данных разбиваются на равные фрагменты (батчи). Каждый батч

имеет одинаковый размер (5000, 10000, 15000 и 20000 образцов). *k* образцов случайным образом выбираются в батче для инициализации центроидов. В исследовании предложенный подход сравнивался с алгоритмом k-средних при количестве кластеров равным 2, 3, 5, 10 и 15, чтобы доказать его эффективность в параллельной кластеризации больших данных. Сравнение времени выполнения предложенного подхода и k-средних для двух реальных наборов данных показано на рис. 3.

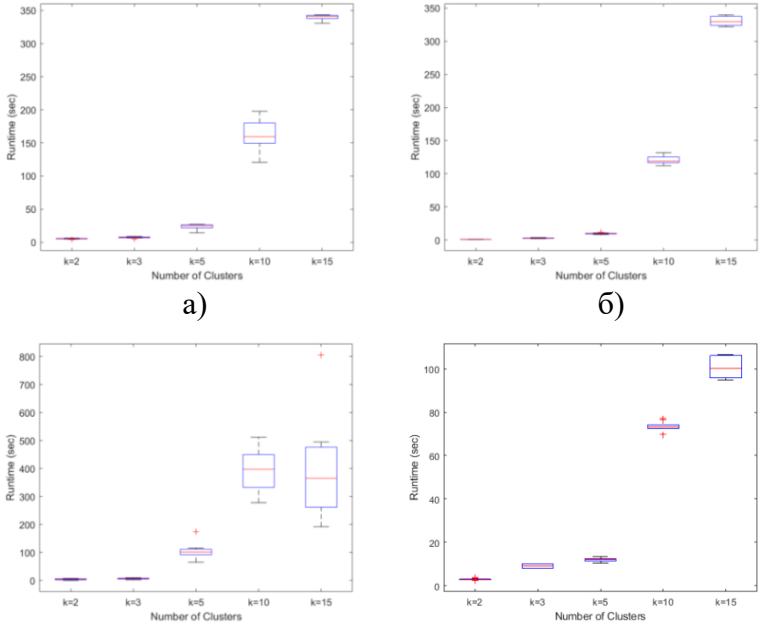


Рис. 3. Сравнение производительности предложенного метода и алгоритма k-средних на наборах данных Phone Accelerometer (а, б) и Individual Household Electric Power Consumption (в, г)

Для дополнительной оценки предложенного метода был проведен анализ при различном количестве батчей на примере набора данных Individual Household Electric Power Consumption

(Рис. 4). При k=2 и k=3 предложенный метод мало зависит от изменения количества батчей. Предлагаемый подход обеспечивает более заметное улучшение с увеличением количества кластеров и батчей. Таким образом, можно увеличить количество батчей (>200), что и было сделано ранее.

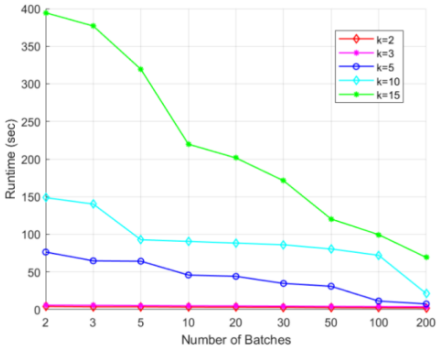


Рис. 4. Производительность предложенного метода при различном количестве батчей для набора данных Individual Household Electric Power Consumption

На основании результатов экспериментов была доказана вычислительная эффективность метода по сравнению с алгоритмом k-средних. Анализируя влияние размера батча на среднее значение функции и среднее время выполнения метода, можно сделать вывод, что эффективность достигается даже при небольшом размере батча. Несмотря на увеличение количества кластеров, скорость предложенного подхода значительно выше по сравнению с алгоритмом k-средних. Метод способен кластеризовать большие данные КФС за короткое время и может применяться для анализа больших данных с целью снижения риска выхода системы из строя. Разбиение данных на небольшие блоки увеличило скорость сходимости и снизило вычислительные затраты алгоритма кластеризации. Метод сравнивался с алгоритмом k-средних, чтобы доказать его эффективность в параллельной кластеризации больших данных КФС. Метод значительно сократил время кластеризации по сравнению с алгорит-

мом k-средних.

В четвертом параграфе представлен метод кластерного анализа больших данных КФС на основе консенсусного ансамбля. Консенсусный подход заключается в поиске согласованного решения за счет возможности совместного использования методов кластерного анализа. Построить наиболее подходящую схему кластеризации для конкретной области можно, применив консенсусный подход к разным наборам алгоритмов в соответствии с их преимуществами и отличительными особенностями. При выработке окончательного решения рассматриваются разные точки зрения, которые не только не противоречат, но, наоборот, компенсируют недостатки каждого метода. Консенсусная кластеризация помогает создавать «надежные» разбиения, обрабатывать шумы и выбросы. Она действует как многообещающее решение для кластеризации больших данных КФС систем. В данном случае проблемой является разработка функции полезности и ее эффективная оптимизация.

Цель исследования состоит в том, чтобы найти консенсусное разбиение π путем решения следующей оптимизационной задачи:

$$\pi^* = \underset{\pi}{\operatorname{argmax}}(w_i U(\pi, \pi_i)), \tag{14}$$

где π^* является функцией консенсуса, U представляет функцию полезности, которая измеряет сходство между π и любым π_i, а разбиение π_i является базовым разбиением $(1 \le i \le r)$ и $w_i \in [0,1]$, $\sum_{i=1}^{r} w_i = 1$.

Веса назначаются отдельным методам кластеризации с использованием функции полезности на основе purity (чистота):

$$f = (1-\lambda) \cdot \sum_{i=1}^{r} w_i U(\pi, \pi_i) + \lambda \cdot \|w\|^2 \longrightarrow \max \tag{15}$$

при условии

$$\sum_{i=1}^{r} w_i = 1, w_i \ge 0, \forall i. \tag{16}$$

где $0 \le \lambda \le 1$ – параметр регуляризации, который указывает на компромисс между максимизацией взвешенной функции полез-

ности и гладкостью, обеспечиваемой w.

Статистический анализ кластеризации больших данных демонстрирует возможность применения взвешенной консенсусной кластеризации к анализу больших данных КФС без меток. С этой целью функция полезности, основанная на Purity, была заменена функцией, основанной на индексах Davies-Bouldin и Calinski-Harabasz. Индекс Davies-Bouldin (DB) основан на соотношении расстояний внутри кластера и между кластерами. DB рассчитывается следующим образом:

$$DB = \frac{1}{k}\sum_{i=1}^{k} R_i \tag{17}$$

$$R_i = \max_{i \neq j}\left(\frac{\delta(C_i)+\delta(C_j)}{dist(C_i,C_j)}\right), \tag{18}$$

где k – количество кластеров, δ – дисперсия внутри кластера, $dist$ – расстояние между i и j кластерами. Целевое значение — это минимум индекса.

Индекс Calinski-Harabasz (CH) характеризуется следующей функцией:

$$CH = \frac{B(k)/(k-1)}{W(k)/(n-k)}, \tag{19}$$

где

$$B(k) = \sum_{i=1}^{k} n_i\, dist^2(O_i, O), \tag{20}$$

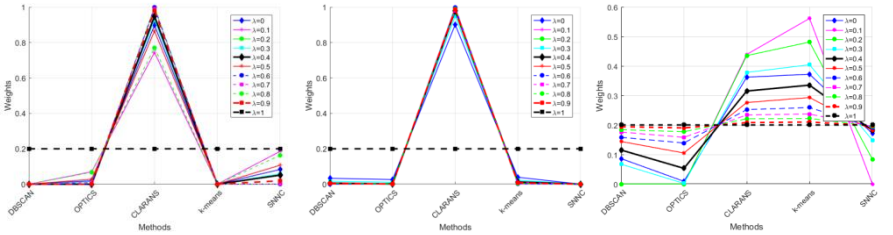
$$W(k) = \sum_{i=1}^{k} \sum_{x \in C_i} dist^2(x, O_i), \tag{21}$$

k – количество кластеров, n – количество объектов в рассматриваемом наборе данных D, C_i – i-ый кластер, n_i – количество объектов в C_i, O – центр набора данных D, O_i – центр C_i в наборе данных, $W(k)$ – сумма внутрикластерных дисперсий для всех кластеров и $B(k)$ представляет собой взвешенную сумму квадратов расстояний между C_i и набором данных D. Наиболее вероятным числом кластеров является значение k, при котором индекс CH достигает своего максимального значения

Экспериментальные результаты на наборах данных различной размерности с использованием метрик расстояния показали, что предложенный метод является высокоэффективным и превосходит современные методы по качеству кластеризации данных. На рис. 5 можно увидеть влияние параметра λ на веса пяти

методов кластеризации для рассмотренных наборов данных NSL-KDD, Phishing и Phone Accelerometer. Результаты экспериментов доказали эффективность предложенного метода к кластеризации данных по сравнению с методами DBSCAN (Density-Based Spatial Clustering of Applications with Noise), OPTICS (Ordering Points to Identification the Clustering Structure), CLARANS (Clustering Large Applications with Randomized Search), k-средних и SNNC (Shared Nearest Neighbor Clustering). Для набора данных Phishing, предложенный консенсусный подход с квадратом евклидова расстояния дал наилучший результат согласно метрикам Purity (0.7166), Mirkin (0.4062), F-мера (0.7283) и PC (0.3053). Результаты основных разбиений превзошли показатель VI и составили 0.0948. Предложенный метод с использованием квадрата евклидова расстояния показал наилучший результат по всем пяти метрикам и совпал с результатом метода CLARANS для набора данных NSL-KDD. А для набора данных Phone Accelerometer предложенный подход превзошел такие методы кластеризации, как DBSCAN и OPTICS по Purity, Mirkin, F-мере и PC, но немного уступил по VI.

Оценка предложенного метода при использовании различных метрик расстояния показала, что метод является высокоэффективным и превосходит современные подходы по качеству кластеризации данных КФС.



а) NSL KDD б) Phishing в) Phone Accelerometer

Рис. 5. Зависимость весов методов кластеризации от параметра λ

Таким образом, проведенное сравнение с известными методами (DBSCAN, OPTICS, CLARANS, k-средних и SNNC) показало,

что наилучший результат получен при применении предложенного метода с использованием квадрата евклидовой метрики.

Третья глава посвящена разработке методов и алгоритмов обеспечения кибербезопасности информационных технологий КФС. Кибератаки являются одной из причин аномальных явлений, наблюдаемых в работе ИТ КФС, а также при передаче трафика по сети. Аномалии сетевого трафика могут привести к некорректной работе отдельного канала или целых сегментов сети, что приведет к отказу в обслуживании на данном сетевом оборудовании. Сетевые атаки постоянно меняются, поскольку злоумышленники используют индивидуальные подходы. На это также влияют изменения в программном и аппаратном обеспечении. Глава состоит из трех параграфов.

В первом параграфе представлен метод классификации кибератак на КФС с применением экстремального машинного обучения (Extreme Learning Machine, ELM). Для выбора наиболее информативных признаков был применен генетический алгоритм поведения «светлячков» с целью повышения точности и увеличения скорости работы алгоритма классификации. Метод ELM имеет высокую скорость вычислений и не требует итеративной настройки параметров для обучения.

Эксперименты проводились на наборе данных NSL-KDD. Все атаки в NSL-KDD поделены на четыре группы: атака «отказ в обслуживании» (Denial of Service Attack, DoS), атаки пользователей на корневой компьютер (Users to Root Attack, U2R), удаленные атаки на локальный компьютер (Remote to Local Attack, R2L) и атаки зондирования (Probe). Каждая запись имеет 41 признак. Метод оценивается при различных функциях активации: функция активации радиального базиса (radial basis activation function, RBF), функция активации треугольного базиса (triangular basis activation function, TriBas) и функция активации Гаусса (Gaussian) (Таблица 2). Применялась функция активации Гаусса для различных значений параметров дисперсии (σ) и математического ожидания (μ). Наилучшая точность обнаружения DoS-, U2R- и R2L-атак была достигнута для функции ак-

тивации Гаусса при $\sigma = 0.1$ и $\mu = 4$ (DoS – 86.89%, U2R – 99.99% и R2L – 99.94%). В случае Probe-атак высокая точность (99.06%) была получена для функции активации Гаусса при $\sigma = 0.2$ и $\mu = 0$.

Метод сравнивался с алгоритмами машин опорных векторов (support vector machines, SVM), случайного леса (random forest, RF), искусственной нейронной сети и глубокой нейронной сети и продемонстрировал наиболее высокую производительность (Таблица 3). Таким образом, предложенный метод при использовании функции активации Гаусса продемонстрировал высокие результаты, как по скорости, так и по точности классификации.

Таблица 2

Сравнение точности обнаружения вторжения при различных функциях активации

Функция активации	DoS	Probe	U2R	R2L
RBF	80.01%	91.79%	99.88%	**99.94%**
TriBas	81.60%	95.62%	99.65%	99.57%
Gaussian (σ =0.2, μ =0)	84.26%	**99.06%**	99.95%	99.70%
Gaussian (σ =0.1, μ =4)	**86.89%**	90.01%	**99.99%**	**99.94%**

Рассмотренный классификатор на основе метода ELM обеспечивает приемлемое качество классификации кибератак и высокую производительность, что делает его привлекательным решением для систем обнаружения вторжений в ИТ КФС в режиме реального времени.

Таблица 3

Оценка производительности предлагаемого подхода для набора данных NSL-KDD

Метод	Bayes Net	Logit Boost	IBk	SVM	RF	DNN	MLP	ELM (Gauss)
Accuracy (%)	69.9	78.8	99.6	93.8	97.93	91.5	81.43	**99.78**

Предложенный метод при сравнении с алгоритмами машин опорных векторов, случайного леса, искусственной нейронной сети и глубокой нейронной сети показал свою эффективность и продемонстрировал наиболее высокую точность при использовании функции активации Гаусса (99.78%) (Таблица 3).

Во втором параграфе представлен метод обнаружения DoS атак на КФС с применением ансамбля классификаторов. В результате, DoS-атаки на сервер или сеть передачи данных ограничивают услуги, а это, в свою очередь, может привести к отказу КФС.

Предложенный метод, указывающий вероятность принадлежности к определенным классам, возвращает вектор оценок классификации для каждой точки данных. Особенность предложенного подхода состоит в том, что для каждой точки из набора данных полученная метка класса соответствует максимальному значению среди всех оценок, полученных методами классификации для данной точки [12].

В качестве классификаторов были рассмотрены деревья решений, алгоритм k-ближайших соседей (K-nearest neighbors, KNN), машины опорных векторов с различными функциями ядра и наивный байесовский классификатор (Naïve Bayes classifier, NB) [10]. Наиболее точный результат показал ансамбль из пяти классификаторов (Таблица 4). Несмотря на то, что NB показывает самый низкий результат (80.45%), при добавлении его в ансамбль классификаторов точность предложенного подхода увеличилась и составила 92.33% для пяти классификаторов [10].

На основании полученных рангов для метрик purity, Mirkin, F-мера, VI и PC был рассчитан результирующий ранг для всех методов кластеризации:

$$rank(method) = \sum_{s=1}^{M} \frac{(M-s+1)\cdot r_s}{M}, \qquad (22)$$

где M – число методов, r_s - количество раз, когда метод появляется в s ранге.

Ансамбль DT+KNN+SVM(Polynom)+NB+SVM(Linear) показал лучший ранг по всем четырем метрикам accuracy, precision,

recall и F-мера для класса DoS.

Таким образом, эксперименты на различных классификаторах доказали превосходство предложенного метода при применении ансамбля из пяти классификаторов [12].

Таблица 4

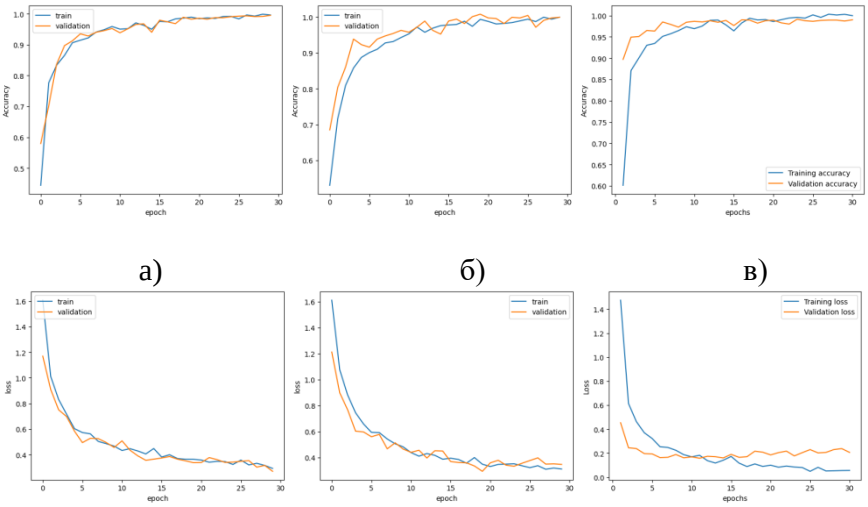
Сравнение предложенного метода с другими классификаторами по метрике accuracy [10, 12]

Класс / Метод	DoS	«Нормальное» состояние	Другие атаки
DT	86.32% [7]	77.19% [7]	64.25% [8]
KNN	88.21% [3]	79.67% [3]	65.80% [6]
SVM(Linear)	87.21% [5]	77.96% [6]	66.22% [5]
SVM(Polynom)	86.64% [6]	79.50% [4]	68.31% [3]
SVM(RBF)	87.25% [4]	78.84% [5]	65.38% [7]
NB	80.45% [8]	71.25% [8]	66.50% [4]
DT+KNN+ SVM(Polynom)	90.74% [2]	84.77% [2]	74.53% [2]
DT+KNN+ SVM(Polynom)+NB +SVM (Linear)	**92.33% [1]**	**88.58% [1]**	**83.35% [1]**

В третьем параграфе представлен алгоритм обнаружения вредоносного программного обеспечения в КФС на основе изображений. Существующие технологии позволяют снизить риск заражения вредоносным ПО с помощью различных инструментов. Однако идеальной защиты от все более изощренных типов вредоносного ПО не существует. Разработан алгоритм на основе трансферного обучения, который объединяет визуализацию вредоносного ПО и преобразование Радона. Особенность алгоритма состоит в том, что он не требует разработки новых признаков. В данном исследовании изображения были получены на основе бинарных образцов вредоносного ПО. Преобразование Радона является обратимым преобразованием изображения. В связи с этим оно может рассматриваться как метод представления текстуры изображения, применяется к полученным изображениям

вредоносных программ в оттенках серого. В отличие от большинства методов, преобразование Радона не требует разработки новых признаков и основано на визуальном анализе образцов вредоносного ПО.

Экспериментальные результаты показали, что объединение признаков из двух глубоких нейронных сетей (AlexNet и MobileNet) может эффективно классифицировать вредоносное ПО в КФС даже при небольших изменениях изображения. Кривые потерь при обнаружении и кривые точности классификации изображений из наборов данных Microsoft malware BIG, IoT_Malware и MaLNet-Image показаны на рис. 6.



а) б) в)

г) д) е)

Рис. 6. Кривые точности обнаружения вредоносных программ и потерь для наборов данных вредоносных программ Microsoft malware BIG (а,г), IoT_Malware dataset (б,д) и MaLNet-Image (в,е)

По данным метрики F-мера для набора данных Microsoft malware BIG классы вредоносного ПО Lollipop, Vundo, Kelihos_ver3, Kelihos_ver1 и Obfuscator.ACY показали результа-

ты >90%. Однако семейства вредоносных программ Ramnit, Tracur и Gatak были менее точно распознаны по показателю precision и составили 76%, 82% и 90%, соответственно. Для набора данных MaLNeT+Images, согласно метрикам precision, recall и F-мера, классы вредоносного ПО Addisplay и SPR были распознаны с 100% точностью. При этом предложенный подход также показал достаточно высокую точность для других рассмотренных классов вредоносного ПО. Для четырех классов набора данных IoT_Malware по метрике precision были получены следующие результаты: Benign (99.50%), Gafgyt (97.99%), Tsunami (96.09%) и Mirai (97.88%). Наиболее точно по метрике recall был идентифицирован класс семейства вредоносных программ Mirai, нацеленный на заражение КФС. В то же время метрики precision и F-мера позволяют определить класс Benign почти с 100% точностью.

Сравнение с различными глубокими нейронными моделями Xception, Inception, EfficientNet, LeNet и другими доказало превосходство предложенного алгоритма (Таблица 5).

Таблица 5

Оценка эффективности предложенного подхода с использованием методов на основе глубокого обучения

Метод	Набор данных	Точность (%)
VGG16	Microsoft malware dataset	98.94
MobileNet		99.25
Xception		99.17
LeNet, AlexNet, InceptionV3		99.70
CNN		98.64
InceptionV3		99.60
Предложенный подход		**99.89**
Adversarial learning	IoT_Malware dataset	97.67
CNN		95.00
Предложенный подход		**99.95**
Vision Transformer	MalNet-Image dataset	97.00
EfficientNetB0+SVM+RF		92.90
Предложенный подход		**99.20**

Исследование продемонстрировало стабильную производительность с точностью 99.89%, 99.95% и 99.20% для наборов данных вредоносных программ Microsoft malware BIG, IoT_Malware и MalNeT-Images соответственно.

Четвертая глава посвящена разработке методов обеспечения защищенности ОТ КФС. Аномальное состояние КФС может быть вызвано неисправными компонентами, временными отказами, неправильной настройкой, кибератаками или их сочетанием. Злоумышленник вмешивается в КФС, чтобы манипулировать показаниями сенсоров или исполнительных механизмов, что приводит к нештатной работе системы. Обнаружение аномалий в промышленном сценарии имеет важное значение, поскольку необнаруженные сбои могут привести к критическим повреждениям. Раннее обнаружение аномалий может повысить надежность подверженного сбоям промышленного оборудования и снизить затраты на эксплуатацию и техническое обслуживание. Глава состоит из четырех параграфов.

В первом параграфе представлен метод обнаружения аномалий в ОТ КФС на основе иерархических скрытых Марковских моделей (ИСММ). В этом случае наблюдаемые события моделируются с помощью СММ. ИСММ подходит для задач со сложной иерархической структурой и зависящих от времени. Нормальные и аномальные состояния, полученные из модели ИСММ можно использовать в качестве входных данных для идентификации атак на КФС (Рис. 7).

Первым шагом в обнаружении аномалий КФС с использованием ИСММ является сбор данных сенсоров. Затем среди них выбираются наиболее информативные признаки. Использование методов векторного квантования (например, алгоритма k-средних) создает значимые наблюдения для модели. Алгоритм Баума-Уэлча выполняет коррекцию параметров модели на этапе обучения. Вероятная последовательность состояний СММ на первом уровне приводит к генерации вектора признаков. Применение

векторного квантования к этим признакам создает последовательность наблюдений из СММ второго уровня и т. д. Затем на этапе тестирования определяется последовательность состояний с помощью алгоритма Витерби, который использует параметры этапа обучения.

Если не обнаружено «необычное» поведение в компонентах КФС, то модель переходит на следующий иерархический уровень и производит проверку на попытку кибератаки на систему.

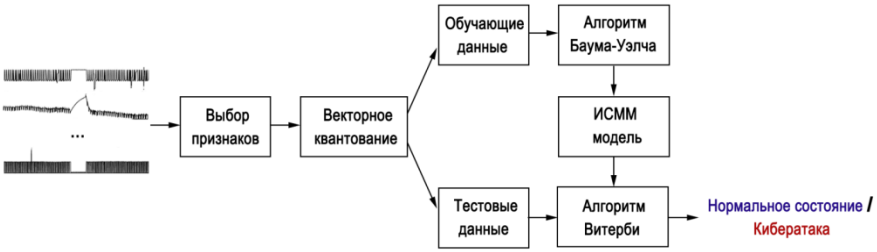


Рис. 7. Общая схема предложенного метода

Эксперименты проводились на наборе данных системы водоочистки. Метод позволил выявить нормальное и аномальное состояния КФС. Он требует значительно меньше обучающих данных для достижения высокой производительности.

Было проведено сравнение с такими моделями, как глубокая нейронная сеть, машины опорных векторов, TABOR (Time Automata and Bayesian netwORk), одномерная сверточная нейронная сеть, LSTM-CUSUM и MAD-GAN (Таблица 6).
Оценка предложенного метода с использованием метрики F-мера дает более высокие показатели. Эксперименты доказали успешность обнаружения аномалий в рассмотренной КФС с помощью разработанного метода.

Преимущество, которое можно ожидать от использования ИСММ по сравнению с большой одноуровневой СММ моделью, заключается в том, что ИСММ с меньшей вероятностью пострадает от переобучения, поскольку отдельные подкомпоненты обучаются независимо на меньших объемах

данных.

Таблица 6

Оценка предлагаемого подхода

Метод	Метрики		
	Precision	Recall	F-мера
DNN	0.9830	0.6785	0.8028
SVM	0.9250	0.6990	0.7963
TABOR	0.8617	0.7880	0.8232
1D CNN	**1.0000**	0.8530	0.9200
LSTM-CUSUM	0.9070	0.6770	0.7750
MAD-GAN	0.9610	**0.9420**	0.9510
Предложенный метод	0.9998	0.9164	**0.9563**

Следствием этого является то, что ИСММ модель требует значительно меньше обучающих данных для достижения производительности, сравнимой с СММ.

Во втором параграфе представлен метод обнаружения аномалий в акустических сигналах КФС с применением трансферного обучения. Архитектура глубокого обучения используется для работы с акустическими сигналами полученными от сенсоров устройств КФС, которые предварительно обрабатываются для получения спектрограмм и скалограмм. Рассматривается функция активации SPOCU (scaled polynomial constant unit) для повышения точности предлагаемого метода. Алгоритм экстремального градиентного бустинга (eXtreme Gradient Boosting, XGBoost) был использован, поскольку он обладает высокой производительностью и требует небольших вычислительных ресурсов на этапе обучения [26].

Для проведения экспериментов был применен набор данных содержащий акустические сигналы четырех различных типов устройств КФС: клапанов, насосов, вентиляторов и направляющих. Для каждого типа устройств рассматривались различные реальные аномальные сценарии: загрязнение, утечка, дисбаланс вращения, повреждение рельса и т. д. Для извлечения признаков из изображений спектрограмм и скалограмм рассматриваются следующие предварительно обученные

глубокие нейронные сети: Xception, MobileNet, DenseNet и Inception. Для оценки производительности метода было проведено его сравнение с глубокими нейронными моделями для различных типов устройств. Применение различных глубоких нейронных моделей показало, что Densenet+XGBoost превосходит другие рассмотренные модели в обнаружении аномалий из сигналов оборудования по метрике F-мера (Таблица 7).

Предложенная модель позволила добиться значительного улучшения обнаружения аномалий по данным сенсоров устройств КФС согласно AUC (area under the ROC curve) 95.45% по сравнению с ранее предложенными моделями.

Таблица 7

Оценка эффективности предложенного метода [26]

Модель	Метрика	Устройство			
		Вентилятор	Насос	Направляющие	Клапан
Inception+ XGBoost	Recall	87.0	95.0	98.0	100
	Precision	89.1	87.9	100	94.3
	F-мера	92.3	90.6	98.1	96.3
Xception+ XGBoost	Recall	100	88.0	98.0	100
	Precision	84.2	100	100	92.6
	F-мера	96.8	92.8	98.1	95.3
Mobilenet +XGBoost	Recall	96.0	88.0	94.0	100
	Precision	78.7	95.7	100	96.2
	F-мера	92.0	90.9	96.1	97.2
Densenet+ XGBoost	Recall	99.9	96.0	98.0	100
	Precision	87.0	100	100	97.1
	F-мера	98.2	97.1	98.1	97.7

В третьем параграфе представлен метод обнаружения атак на устройства КФС с применением глубокой гибридной модели. Предложенный метод сочетает в себе преимущества одномерной конволюционной нейронной сети (convolutional neural network, CNN), управляемой рекуррентной нейронной сети (gated recurrent unit, GRU) и нейронной сети с долгой

краткосрочной памятью (long short-term memory, LSTM). Выходные данные моделей GRU, CNN и LSTM объединяются для повышения точности обнаружения кибератак на КФС и передаются на два полносвязных слоя, состоящих из 150 нейронов. За ними следует слой softmax. Для повышения точности метода рассматривается функция активации SPOCU. Оптимизатор AdamW+Amsgrad применяется для уменьшения ошибки обучения глубокой нейронной модели и увеличения скорости обучения Эксперименты проводились на двух наборах данных: SWaT (secure water treatment) и GHL (gasoil heating loop), которые содержат информацию о «нормальном» состоянии КФС и сбоях, вызванных кибератаками (Рис. 8).

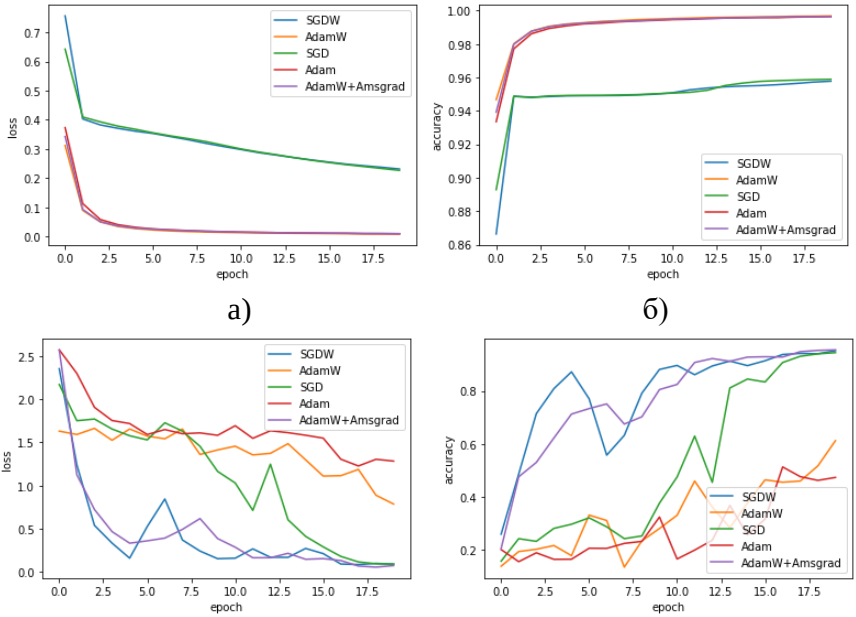


Рис. 8. Кривые потерь и точности обнаружения кибератак для предложенного метода на наборах данных SWaT (а,б) и GHL (в,г)

Метод оценивался с использованием различных оптимизаторов: стохастический градиентный спуск (SGD), SGD с сокращением весов (SGDW), Adam, Adam с сокращением весов (Adam with a weight decay, AdamW) и AdamW+Amsgrad. Наиучший результат показал предложенный метод с оптимизатором AdamW+Amsgrad (Рис. 8).

Предложенный метод оценивался с применением метрик precison, recall и F-мера. Эксперименты показали, что предложенный метод обеспечивает высокую точность обнаружения кибератак на рассмотренные КФС по сравнению с известными методами машинного обучения, такими как простая глубокая нейронная сеть, машины опорных векторов, LSTM, GRU и CNN (Таблица 8).

Таблица 8
Оценка предложенного метода на различных наборах данных

Набор данных	Метрики		
	Precision (%)	Recall (%)	F-мера (%)
Secure Water Treatment dataset	99.76	99.75	99.74
Gasoil Heating Loop dataset	97.06	95.66	96.04

В четвертом параграфе представлен метод классификации отказов КФС на основе изображений. Оценка производительности и своевременное обнаружение неисправностей КФС могут снизить эксплуатационные расходы и расходы на техническое обслуживание. Разработанный метод сочетает в себе преимущества глубокой нейронной сети (deep neural network, DNN) и CNN. Частотные и временные признаки, а также изображения вибрационного сигнала рассматриваются как входные данные для глубокой гибридной модели. Изображения спектрограммы кратковременного преобразования Фурье и скалограммы непрерывного вейвлет-преобразования

сигналов сенсоров КФС считаются входными данными для предложенного метода.

В качестве примера была рассмотрена задача обнаружения неисправностей нефтяного электрического погружного насоса (ЭПН). Исследуемый ЭПН состоит из шести компонентов: двух двигателей, двух протекторов и двух насосов. Вибрационные сигналы генерируются во время работы системы. Они несут в себе наиболее важную информацию о состоянии механических устройств, включая нефтяное оборудование. Характеризующие неисправности компонентов КФС признаки извлекаются для интеллектуального анализа необработанных сигналов и повышения точности диагностики.

Анализ графического представления ROC-кривой позволяет оценить качество предложенной глубокой гибридной модели (Рис. 9).

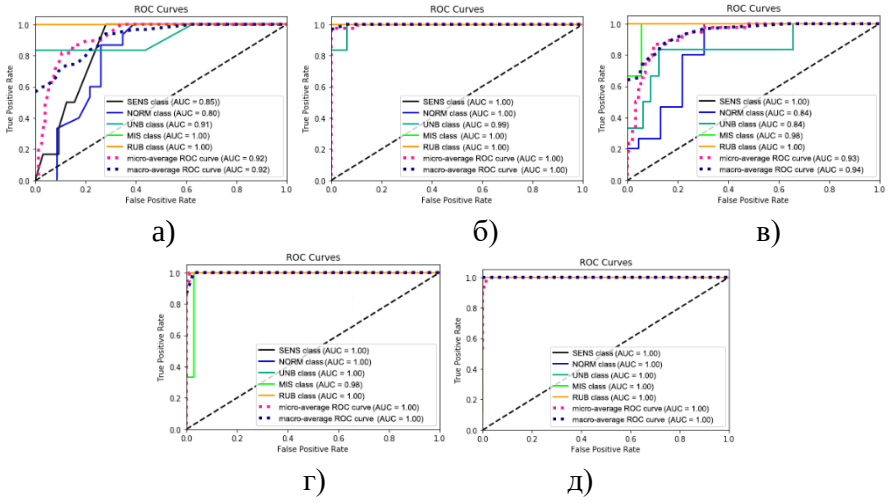


Рис. 9. ROC-кривые оценки точности классификации рассматриваемых методов для пяти классов: faulty sensor (SENS), normal operational condition (NORM), unbalance (UNB), misalignment (MIS) и rubbing (RUB) [20]

Экспериментальные результаты показывают, что это лучшая

модель для выявления неисправностей ЭПН по рассмотренным признакам, полученным из сигналов вибрации (площадь под ROC-кривой) составила 100% (Рис. 9).

Экспериментальная оценка показывает, что предложенная глубокая гибридная модель превосходит ранее предложенные методы на основе глубокого обучения, в том числе k-ближайших соседей (KNN) согласно метрике recall (Таблица 9).

Таблица 9

Сравнение эффективности предлагаемого подхода с другими методами на основе метрики recall

Метод	Классы				
	NORM (%)	UNB (%)	MIS (%)	RUB (%)	SENS (%)
Ансамблевый подход	97.82	97.51	76.00	48.29	95.65
KNN	97.50	89.50	60.00	33.50	87.00
KNN+	97.50	89.50	68.00	38.00	89.00
KNN+FS	97.50	90.50	71.00	35.50	89.50
KNN+EFS	98.00	93.00	81.00	34.50	91.00
Предложенный метод	**100**	**100**	**99.93**	**100**	**100**

Результаты этого исследования показывают, что предложенная глубокая гибридная модель может автоматически и одновременно извлекать характеристики сигналов сенсоров, которые чувствительны к сбоям вызванным кибератакми во временной, частотной и частотно-временной областях. Экспериментальная оценка показала, что предложенный метод превзошел известные методы машинного обучения при классификации отказов рассмотренной КФС. Таким образом, предложенная гибридная модель с использованием глубоких нейронных сетей может быть применена при диагностике отказов оборудования КФС.

Пятая глава посвящена разработке методов оценки киберустойчивости КФС. Наиболее важной задачей является сохранение способности КФС корректно функционировать даже в усло-

виях деструктивных информационных воздействий, поскольку успешное осуществление кибератак на такие системы может привести к негативным финансовым последствиям и экологическим катастрофам, а также привести к гибели людей [30]. Поскольку злоумышленники часто ищут и нацеливаются на самое слабое звено – наиболее уязвимый функциональный компонент в КФС – самые уязвимые устройства могут стать хорошей отправной точкой для эффективного исследования киберустойчивости системы. Существующие методы не могут всесторонне анализировать распространенные многоэтапные кибератаки в среде КФС [30]. Глава состоит из двух параграфов.

В первом параграфе представлен метод определения критичности уязвимостей функциональных компонентов КФС систем на основе Байесовского графа атак. Интеграция интеллектуальных устройств и технологий в различных отраслях существенно изменила всю технологическую инфраструктуру. КФС распознает окружающую среду с помощью сенсоров, принимает решения в соответствии со своим назначением и обеспечивает корректирующие действия с помощью исполнительных механизмов. Связь между сетями ИТ и ОТ важна при моделировании КФС систем. Конвергенция ИТ и ОТ растет, что позволяет более эффективно управлять и эксплуатировать КФС. Серьезность отказа компонента КФС влияет на систему в зависимости от количества компонентов и их разнообразия. Это доказывает необходимость измерения критичности уязвимостей компонентов системы. Своевременное выявление критических недостатков безопасности в КФС позволяет обнаруживать риски и потенциальные угрозы. Для решения этой проблемы создаются модели угроз, позволяющие лучше понять потенциальные уязвимости, которые необходимо учитывать для обеспечения надежности системы. Выбор оптимального решения для оценки критичности функциональных уязвимостей компонентов КФС — сложный процесс, поскольку все уязвимости должны быть идентифицированы, классифицированы и количественно оценены в соответствии с единым подходом в рамках процесса кибе-

рустойчивости. Статистика за последние несколько лет показывает рост количества кибератак на КФС, при этом в большинстве случаев целью злоумышленников является получение контроля над подсистемой управления [30].

Однако решение этой проблемы возможно с использованием байесовского графа атак (БГА), который позволяет оценить критичность уязвимостей компонентов КФС. Он предоставляет информацию о связях между уязвимостями, что является важным фактором. БГА зависит от операционных систем и сетевых протоколов, правил контроля доступа КФС, идентификации программного обеспечения и информации об уязвимостях.

Методы многокритериального принятия решений могут быть использованы в качестве подходящего подхода для решения задачи ранжирования критичности уязвимостей компонентов КФС. Они помогают лицам, принимающим решения, выбирать оптимальные альтернативы на основе заданных критериев. Для определения критичности уязвимостей компонентов КФС предлагается использовать многокритериальный метод принятия решений Promethee II. Он считается довольно простым среди методов многокритериального анализа и все чаще используется экспертами, принимающими решения. Метод дает довольно стабильные результаты и основан на попарном сравнении альтернатив, соответствующих каждому критерию. Promethee II позволяет упорядочить и выявить наиболее критичные уязвимости рассматриваемой системы. Уязвимость количественно оценивается с использованием Общей системы оценок уязвимостей (Common Vulnerability Scoring System, CVSS) и Национальной базы данных уязвимостей (National Vulnerability Database, NVD).

Для иллюстрации потенциальных сценариев кибератак, использующих уязвимости в компонентах КФС, можно рассмотреть пример транспортировки природного газа как одного из критически важных секторов инфраструктуры (Рис. 10). Природный газ транспортируется по трубопроводам компрессорными станциями. Они состоят из нескольких газовых компрессоров, которые перемещают газ по трубопроводу. После фильтра-

ционной сепарации природный газ подается в компрессорную установку. Системы смазки и охлаждения двигателя компрессора служат для защиты и поддержания скорости потока газа. ОТ рассматриваемой КФС представляют собой программируемые логические контроллеры (programmable logic controller, PLC) и удаленные терминальные устройства (remote terminal units, RTU), подключающиеся к сенсорам и исполнительным механизмам. Сенсоры передают сигналы на контроллеры, которые отправляют управляющие сигналы на исполнительные механизмы. Операторы процесса взаимодействуют с системой управления через человеко-машинный интерфейс (Human-Machine Interface, HMI). Архив данных хранит значения параметров процесса. Сервер данных получает информацию от контроллера и передает ее другим контроллерам или HMI. Сеть управления передает инструкции и данные между блоками управления и измерения и устройствами SCADA. Инженерные рабочие станции могут быть подключены к корпоративной сети или Интернету. Здесь устанавливаются средства разработки программного обеспечения. Компрессорные станции при транспортировке природного газа обычно контролируются с помощью распределенной системы управления, возможно, сопряженного с отдельной системой безопасности. Станции управляются удаленно через устройства RTU, подключенные к центральной системе SCADA (Supervisory Control and Data Acquisition system) WAN (Wide Area Network).

Предполагается, что целью злоумышленника может быть компрометация архива данных или рабочей станции (Engineering Workstation, EWS), которые обычно являются основными целями из-за их взаимосвязанного характера. Исследуются три сценария кибератак на систему транспортировки природного газа (Рис. 10):

• Манипулирование элементами управления

В зависимости от конфигурации системы злоумышленник, способный взаимодействовать с WAN SCADA, может использовать CVE-2020-13500 или CVE-2022-29966 для манипулирова-

ния системными настройками, файлами или критическими значениями, связанными с компрессорными станциями. Он может использовать CVE-2022-33139 для обхода аутентификации, а затем манипулировать значениями мониторинга, связанными с компрессорными станциями, чтобы перегружать операторов ложными сигналами тревоги или изменять заданные значения расхода, чтобы нарушить транспортировку природного газа, тем самым расширяя возможности манипуляции.

• Отказ в управлении

Злоумышленник, имеющий возможность взаимодействовать с RTU, может использовать уязвимости CVE-2022-29961 или CVE-2022-29955 для обхода аутентификации в RTU и ввода команд, которые останавливают операцию, а, воспользовавшись уязвимостью CVE-2022-30262, он может лишить операторов возможности управлять и контролировать работу компрессорных станций PLC1 и PLC2. Альтернативно, злоумышленник, способный взаимодействовать с сетевым коммутатором (network switcher, NS), может воспользоваться неаутентифицированным соединением (CVE-2021-30276 или CVE-2021-31886) и отправить определенные команды на устройства RTU, что может приостановить работу или нарушить связь между SCADA WAN и RTU, лишая операторов возможности мониторинга и управления компрессорными станциями.

• Потеря контроля

Для этого злоумышленник взламывает систему управления самой компрессорной станции, используя уязвимости CVE-2022-30262, CVE-2022-26657, CVE-30315, CVE-2022-30260 или CVE-2022-31801.

Это позволяет добиться выполнения кода на RTU и без ограничений общаться через различные сетевые интерфейсы. Использование уязвимостей CVE-2022-30313 и CVE-2022-30315 позволяет проникнуть в сеть безопасности. Тем самым злоумышленник производит манипулятивные настройки или даже добивается выполнения кода на программируемом логическом

контроллере PLC2 для отключения систем аварийного отключения и систем пожарной и газовой безопасности.

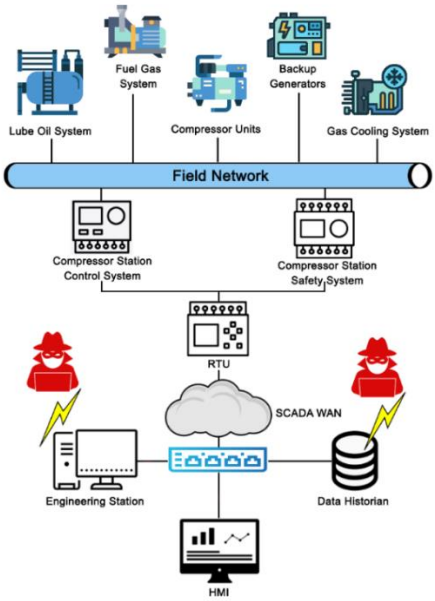


Рис. 10. Пример конфигурации КФС

После того, как устройства ОТ КФС скомпрометированы, злоумышленники могут использовать различные CVSS для запуска атак на процессы управления, которые влияют на физические операции. Информация об уязвимостях взята из описаний NVD. Злоумышленник использует подключение ИТ-сети КФС систем к сети Интернет для выполнения вредоносного кода.

Для проведения экспериментов была рассмотрена информация о различных уязвимостях компонентов КФС, включая базовую оценку, оценку воздействия γ, подоценку возможности использования ε, вектор атаки, взаимодействие с пользователем, требуемые привилегии и оценку сложности эксплойта. Было проведено сравнение предложенного метода с известными многокритериальными методами принятия решений, такими как

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution), VIKOR (VIseKriterijumska Optimizacija I Kompromisno Resenje) и ELECTRE (Elimination and Choice Translating Reality) (Таблица 10).

Таблица 10
Сравнение предложенного метода с известными подходами

Компонент КФС	TOPSIS	VIKOR	ELECTRE	Предложенный метод
PLC_1	4	4	3	4
PLC_2	3	3	4	3
RTU	3	4	3	3
SW	1	1	1	1
DH	2	2	2	2
NS	1	1	1	1
HMI	1	1	1	1
EWS	1	1	1	1

Результаты ранжирования рисков с использованием методов Promethee II и TOPSIS аналогичны. Более того, на третью и четвертую позиции были помещены два разных варианта решения для методов VIKOR и ELECTRE.

Подводя итоги, предложен количественный метод оценки критичности уязвимостей компонентов КФС на основе многокритериального метода принятия решений Promethee II. Promethee II обеспечивает полное ранжирование альтернатив. Он позволяет ранжировать и идентифицировать наиболее критичные уязвимости компонентов рассмотренной системы транспортировки природного газа по трубопроводу. БГА был создан на основе рассматриваемой структуры КФС с учетом известных уязвимостей ее компонентов. В результате удалось выявить наиболее критичные уязвимости функциональных компонентов КФС с помощью Promethee II.

Во втором параграфе представлен метод оценки рисков КФС с применением нечеткого интеграла Сугено. Защита КФС от

кибератак является серьезной проблемой, требующей методов оценки рисков кибербезопасности. В настоящее время методы анализа и оценки рисков КФС можно разделить на качественные и количественные методы. Первые основаны на экспертной оценке для определения вероятности и потенциального воздействия рисков, а также раскрывают природу рисков. При этом количественный анализ рисков использует математические и статистические методы для расчета вероятности и потенциального воздействия рисков. Однако исследователи отдают предпочтение количественным методам анализа и оценки рисков, поскольку они позволяют более точно оптимизировать ресурсы безопасности.

Структура предлагаемой методологии оценки рисков включает моделирование системы, определение и ранжирование критичности компонентов системы и оценку рисков. Смоделированный граф атак, нацеленных на устройства КФС, позволяет получить различные измерения для оценки киберрисков. Различные измерения из кибер- и физической среды объединяются в единую количественную оценку, которая используется для диагностики состояния системы. Значения уязвимостей компонентов в кибер- и физических слоях КФС используются для расчета риска системы на основе нечеткого интеграла Сугено. На основе полученных значений выбираются наиболее критические узлы КФС и прогнозируются возможные пути атаки.

Для обеспечения киберустойчивости КФС на основе графа атак необходимо определить критичность ее узлов. Индексы оценки риска позволяют измерить, насколько уязвима система к кибератакам, и определить расположение компонентов КФС относительно друг друга. Рассматриваемые индексы охватывают следующие две основные области: (1) ОТ и (2) ИТ. Рассматриваемые среды — (1) физическая среда и (2) киберсреда. Показатели ИТ и ОТ прогнозируют состояние КФС при возникновении нежелательного события. Например, для анализа состояния энергетической системы на основе ветра на предмет отказов оператор должен сначала знать такие параметры, как скорость

включения, номинальная скорость, скорость отключения и номинальная мощность на каждом узле системы. Отказ отдельных компонентов КФС может вызвать неконтролируемые отклонения различных параметров из-за кибератак злоумышленников, что в конечном итоге может привести к разрушению системы.

Граф атак позволяет связывать различные уязвимости. Он определяет потенциальные угрозы для узлов рассмотренной КФС системы на основе информации об уязвимостях, чтобы представить пути атаки в системе. Фактически, CVSS оценивает сложность реализации кибератаки, принимая во внимание уязвимости для каждого устройства, которое существует в конкретном узле КФС.

КФС системы имеют разнообразную топологическую структуру. Рассмотрим некоторые индексы для оценки критичности узлов графа атак. Эти индексы предоставляют подробную информацию обо всей сети. Пусть $G = (\mathcal{N}, \mathcal{E})$ – это граф, где $\mathcal{N} = \{1, \dots, \mathrm{n}\}$ – набор узлов (вершин), а $\mathcal{E}$ — набор ребер, $\mathcal{E} = \{(\mathrm{i}, \mathrm{j}) | \mathrm{i}, \mathrm{j} \in \mathcal{N}\}$. n представляет собой общее количество узлов в сети.

а) *Центральность по близости.* Центральность по близости узла измеряет, насколько близко узел (т. е. i) находится ко всем другим узлам, путем вычисления кратчайшей длины пути от одного узла до других узлов в сети. Узлы с высоким показателем центральности по близости имеют большее влияние на другие узлы в сети.

Определение (центральность по близости). Центральность по близости C_i^c для узла i вычисляется следующим образом:

$$C_i^c = \frac{n-1}{\sum_{\substack{j=1 \\ i \neq j}}^{n} d_{ij}}, \; i = 1, \dots, n, \tag{23}$$

где d_{ij} – кратчайшее расстояние от вершины i до вершины j. В графе может быть несколько путей, соединяющих i с j, среди которых d_{ij} – кратчайший путь.

б) *Центральность по собственному вектору* показывает связь между самой «влиятельной» вершиной графа и соседними вершинами.

Определение (центральность по собственному вектору). Центральность по собственному вектору E_i^c для узла i определяется следующим образом:

$$E_i^c = \frac{1}{\delta_{max}} \sum_{j=1}^{n} a_{ij} E_j^c, i = 1, \dots, n, \quad (24)$$

где δ_{max} обозначает наибольшее собственное значение матрицы смежности. Матрица смежности представляет собой квадратную матрицу размера $n \times n$, $\mathbb{A} = \|a_{ij}\|_{i,j=1}^{n}$, элементы которой определяются как: $a_{ij} = 1$, когда узел i соединен с узлом j, и $a_{ij} = 0$ в противном случае.

в) *Центральность по промежуточности ребра* измеряет количество кратчайших путей, проходящих через определенную вершину графа. Эта теоретико-графовая метрика измеряет, как часто узел выполняет роль «моста» на кратчайших путях между двумя другими узлами. При переводе сети в теоретико-графовую модель центральность по промежуточности ребра (B_i^c) для узла указывает на возможность атаки, проходящей через этот узел.

Определение (центральность по промежуточности ребра). Центральность по промежуточности ребра B_i^c узла i определяется следующим образом:

$$B_i^c = \sum_{\substack{k,j=1 \\ k \neq j \neq i}}^{n} \frac{\sigma_{kj,i}}{\sigma_{kj}}, i = 1, \dots, n, \quad (25)$$

где σ_{kj} – общее количество кратчайших путей от исходного узла k до конечного узла j, а $\sigma_{kj,i}$ – количество путей от k до j, проходящих через i.

Кратчайшие пути относятся ко всем кратчайшим путям между каждой парой вершин в графе. Если одна вершина является частью кратчайших путей, то она имеет высокую центральность промежуточности по ребру.

г) *Центральность по Катцу* – это параметр теории графов, который придает важность узлу с учетом структуры сети и положения узла в сети. Он определяет количество узлов, соединенных через этот путь, а вклад удаленных узлов «штрафуется».

Определение (центральность по Катцу). Центральность по Катцу K_i^c узла i определяется как:

$$K_i^c = \sum_{q=1}^{\infty} \sum_{j=1}^{n} \beta^q (\mathbb{A}^q)_{ji}, i = 1, \dots, n, \quad (26)$$

где $\beta \in (0,1)$ – коэффициент затухания, т. е. доля участия удаленных вершин, а $(\mathbb{A}^q)_{ji}$ – общее количество связей степени q между узлами i и j.

Данные индексы предоставляют подробную информацию обо всех КФС имеющих разнообразную топологическую структуру.

Объединение значений критериев оценки рисков КФС выполняется с использованием нечетких интегралов, которые определяются относительно нечетких мер. Предполагается, что значения нечеткой меры и всех входных параметров изменяются в пределах единичного интервала.

Определение (нечеткая мера). Пусть $N = \{1, \dots, n\}$ конечное множество и $\mu: 2^{\mathcal{N}} \to [0,1]$ – это функция такая, что $\mu(\emptyset) = 0$ и $\mu(\mathcal{N}) = 1$. Если для любых A и B, таких, что $A \subseteq B \subseteq \mathcal{N}$, выполняется условие $\mu(A) \leq \mu(B)$, то нечеткое множество μ называется нечеткой мерой.

Определение (λ-мера Сугено). Пусть $N = \{1, \dots, n\}$ конечное множество и $\lambda \in (-1, +\infty)$. Функция $\mu: 2^{\mathcal{N}} \to [0,1]$ является -мерой Сугено, если выполняются следующие условия:

$$\mu(\emptyset) = 0, \quad (27)$$

$$\mu(\mathcal{N}) = 1, \quad (28)$$

$$\mu(A) \leq \mu(B), \forall A, B \text{ таких, что } A \subseteq B \subseteq \mathcal{N}, \quad (29)$$

$$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\, \mu(A)\, \mu(B), \forall A, B \subseteq \mathcal{N} \text{ с } A \cap B = \emptyset, \quad (30)$$

где ∅ — пустое множество.

Уравнения (27) и (28) представляют меры пустого множества и комбинации всех множеств соответственно. Уравнение (29) представляет свойство монотонности. Уравнение (30) представляет возможные подмножества и объединенные подмножества.

При повторном применении (30) для каждого $A \subseteq \mathcal{N}$ значение $\mu(A)$ можно вычислить следующим образом:

$$\mu(A) = \left[\frac{\prod_{i\in A}(1+\lambda\mu(\{i\}))}{\lambda}\right]. \quad (31)$$

Используя ограничение $\mu(\mathcal{N}) = 1$ (28) и применяя (31), значение λ можно вычислить с помощью (32):

$$\lambda + 1 = \prod_{i=1}^{n}(1 + \lambda\mu_i), \quad (32)$$

где $\mu_i = \mu(\{i\})$.

Определение (дискретный интеграл Сугено). Пусть μ – это нечеткая мера в $\mathcal{N}$. Дискретный интеграл Сугено функции $x = (x_1, x_2, \dots, x_n)$: $[0,1]^n \to [0,1]$ относительно μ определяется как:

$$SI_\mu(x) = \max_{1\le i\le n}(min(x_{\pi(i)}, \mu(\{\pi(1), \pi(2), \dots, \pi(n)\}))) =$$

$$= \max_{1\le i\le n}\{min\{x_{\pi(i)}, \mu(\{\pi(1)\})\}, \dots, min\{x_{\pi(n)}, \mu(\{\pi(1), \dots, \pi(n)\})\}\}, (33)$$

где π – перестановка в $\mathcal{N}$ такая, что $x_{\pi(1)} \le \cdots \le x_{\pi(n)}$.

Основная идея интеграла Сугено основана на взвешенном минимуме и максимуме, что позволяет оценить важность каждой модели с использованием нечетких мер. Нечеткий интеграл Сугено определяет наивысший уровень сходства между целевыми и прогнозируемыми значениями. Объединение значений критериев оценки риска КФС выполняется с использованием нечетких интегралов, которые определяются относительно нечетких мер. Предполагается, что значения нечеткой меры и всех входных параметров изменяются в пределах единичного интервала.

В исследовании для оценки рисков путей атаки используется метрика, основанная на нечетком интеграле Сугено. Риск каждого узла графа атак рассчитывается следующим образом:

$$Risk_i = Probability_i \times ImpactSI_i, i = 1, \dots, n, \quad (34)$$

где $Probability_i$ – это вероятность доступа к узлу i, которая показывает количество путей атаки и рассчитывается следующим образом:

$$Probability_i = 1 - \prod_{j=1}^{n}\left(1 - P_j\right),\ i = 1, \dots, n, \quad (35)$$

где P_j – это вероятность атаки на узел j.

Здесь P_i рассчитывается как

$$P_i = AV_i \times AC_i \times UI_i \times PR_i,\ i = 1, \dots, n, \quad (36)$$

где AV_i – вектор атаки, AC_i – сложность атаки, UI_i – взаимодействие с пользователем, PR_i – требование привилегий.

Подверженность риску с использованием нечеткого интеграла Сугено (Sugeno integral, SI) рассчитывается с использованием индексов B_i^c, C_i^c, E_i^c и K_i^c, а также оценок целостности (I_i), доступности (A_i) и конфиденциальности (C_i), полученных из CVSS v3.1, следующим образом:

$$ImpactSI_i = SI(B_i^c, C_i^c, E_i^c, K_i^c, I_i, A_i, C_i),\ i = 1, \dots, n. \quad (37)$$

Узлы с высокими значениями $ImpactSI$ считаются более уязвимыми с точки зрения кибербезопасности КФС. Метрика, основанная на нескольких индексах, лучше характеризует каждый узел системы. В отличие от одного индекса, она более информативна. После расчета требуемых коэффициентов они объединяются с помощью многокритериального анализа решений. Нечеткий интеграл Сугено используется для оценки рисков путей атак для обеспечения киберустойчивости КФС.

Вводится переменная R0 – пороговое значение, при котором узел КФС с Risk выше этого порога считается «неустойчивым». На практике это значение должны определять опытные эксперты. Если узел «неустойчив», немедленно принимаются экстренные меры по устранению риска. Если Risk ниже R0, то узел считается «устойчивым», и риска нет. В этом случае следующая оценка рисков может быть выполнена через определенный интервал. Окончательное решение Risk делает систему более киберустойчивой с точки зрения киберфизической безопасности.

В качестве примера была рассмотрена КФС для генерации энергии ветра (Рис. 11). Энергия, полученная на основе ветра, является одним из критических секторов инфраструктуры. Ветровые турбины широко используются на различных объектах: предприятиях, в домохозяйствах, частных домах и т. д. Потоки ветра вращают лопасти ветряной турбины, приводя ее в движение. Чем сильнее ветер, тем больше вырабатывается энергии. Это вращение запускает турбину, которая также начинает вращаться. Ветровые турбины – это устройства, преобразующие

энергию ветра в электрическую энергию. Энергия передается по валу ротора, который соединен с редуктором, который приводит в действие электрогенератор. Турбина состоит из системы охлаждения, системы мониторинга состояния и флюгера. Они служат входными данными для контроллера, который определяет положение лопастей и ротора. Система управления батареями контролирует и управляет несколькими аккумуляторными батареями и обеспечивает стабилизацию сети. Значения параметров процесса, полученные в ходе мониторинга, хранятся в архиве данных. HMI интерфейс обеспечивает взаимодействие операторов процесса с системой управления. Инженерные рабочие станции содержат средства разработки программного обеспечения. Подключение через RTU связывает ветропарк с центральной системой SCADA. Предположим, что злоумышленник воспользовался уязвимостями компонентов КФС и реализовал следующие сценарии кибератак:

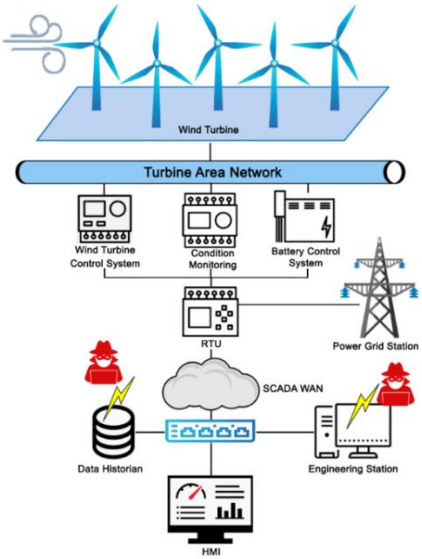


Рис. 11. Примеры сценариев кибератак на КФС

- *Манипуляция и отказ в управлении.* Злоумышленник, способный взаимодействовать с сервером SCADA, может использовать уязвимость CVE-2019-14925 для манипулирования системными конфигурациями, файлами или критическими значениями, связанными с работой ветропарка. Уязвимость может привести к несанкционированному доступу к конфиденциальным данным, включая имена пользователей, пароли и другую конфиденциальную информацию. Злоумышленник также может злоупотребить тем фактом, что соединения не аутентифицированы (CVE-2021-27395), что позволит несанкционированно манипулировать данными и выдавать команды RTU. Это может остановить выполнение логических задач и нарушить связь между SCADA и RTU. Это не позволит операторам контролировать компрессорные станции.
- *Потеря управления.* Чтобы скомпрометировать учетные данные и получить доступ к сети ветропарка, злоумышленник использует уязвимости CVE-2019-9013, CVE-2022-1159 и CVE-2021-22797 для выполнения кода на RTU. Это позволит перейти к управлению отдельными турбинами. Злоумышленник может проникнуть во внутреннюю сеть системы управления, используя уязвимость CVE-2018-5452, чтобы манипулировать конфигурацией системы, рабочими параметрами и прошивкой контроллера (Таблица 11).

Это может привести к отключению функции защиты от превышения скорости, встроенной в RTU, и отключению нагрузки, что приведет к отключению турбины. Используя CVE-2020-7566, злоумышленник может скомпрометировать данные и отключить системы мониторинга работоспособности, которые обеспечивали бы раннее предупреждение об опасности. Кроме того, они могут нацелиться на контроллер PLC и использовать CVE-2020-6992 для компрометации учетных данных и получения кода выполнения на PLC. С помощью этого злоумышленник может повлиять на функции управления батареей, что приведет к простою системы и потенциальной дестабилизации КФС.

Таблица 11

Информация об уязвимостях

Узел	Уязвимость CVE-ID	*BS*	*IS*	ζ	*AV*	*C*	*I*	*A*	*UI*	*PR*	*AC*
1	CVE- 2021-41773	7.5	3.6	3.9	0.85	0.56	0	0	0.85	0.85	0.77
2	CVE-2022-22720	9.8	5.9	3.9	0.85	0.56	0.56	0.56	0.85	0.85	0.77
3	CVE-2022-30522	7.5	3.6	3.9	0.85	0	0	0.56	0.85	0.85	0.77
4	CVE-2014-7844	7.8	5.9	1.8	0.55	0.56	0.56	0.56	0.85	0.85	0.77
5	CVE-2019-9557	6.1	2.7	2.8	0.85	0.22	0.22	0	0.62	0.85	0.77
6	CVE-2020-2512	5.9	3.6	2.2	0.85	0	0	0.56	0.85	0.85	0.44
7	CVE-2020-24673	9.8	5.9	3.9	0.85	0.56	0.56	0.56	0.85	0.85	0.77
8	CVE-2020-6992	6.7	5.9	0.8	0.55	0.56	0.56	0.56	0.85	0.27	0.77
9	CVE-2021-27395	8.1	5.2	2.8	0.85	0	0.56	0.56	0.85	0.62	0.77
10	CVE-2020-3960	8.4	5.8	2.0	0.55	0.56	0	0.56	0.85	0.62	0.77
11	CVE-2021-22797	7.8	5.9	1.8	0.55	0.56	0.56	0.56	0.62	0.85	0.77
12	CVE-2019-14925	6.5	3.6	2.8	0.85	0.56	0	0	0.85	0.62	0.77
13	CVE-2019-9013	8.8	5.9	2.8	0.62	0.56	0.56	0.56	0.85	0.85	0.77
14	CVE-2022-1159	7.2	5.9	1.2	0.85	0.56	0.56	0.56	0.85	0.27	0.77
15	CVE-2020-7566	7.3	5.2	2.1	0.62	0.56	0.56	0	0.62	0.85	0.77
16	CVE-2018-5452	7.5	3.6	3.9	0.85	0	0	0.56	0.85	0.85	0.77
17	CVE-2023-0286	7.4	5.2	2.2	0.85	0.56	0	0.56	0.85	0.85	0.44

С целью проведения экспериментов для каждого узла i системы были рассчитаны значения индексов центральности C_i^c, E_i^c, K_i^c, B_i^c, а также целостности (I_i), доступности (A_i), конфиденциальности (C_i), полученные на основе CVSS. Это позволило выявить критические узлы системы. Индексам были присвоены

«экспертные» веса. Чем выше вес, тем выше «информативность» индекса. В таблице 11 приведены значения индексов для всех узлов рассматриваемой КФС. Эти значения позволяют оценить критичность устройств системы на основе значений воздействия и вероятности.

Предложенный подход был выбран для оценки серьезности уязвимости по критериям:

$$R = \begin{cases} \text{critical}, & v \in [5, 10] \\ \text{high}, & v \in [3, 5) \\ \text{medium}, & v \in [2, 3) \\ \text{low}, & v \in [0, 2) \end{cases}. \quad (38)$$

Были рассмотрены различные пути атак, направленные на возможные узлы графа для изменения их состояний.

В отличие от существующих подходов, данное исследование было направлено на решение конкретной проблемы количественного измерения киберустойчивости КФС систем с учетом факторов, влияющих как на ее физический слой, так и на киберслой. Модель объединяет топологию и уязвимости киберфизической сети, обеспечивая эффективность предложенного метода в сохранении киберустойчивости КФС.

В заключении отражены наиболее важные результаты диссертационной работы, а также сформулированы основные выводы, следующие из предложенных методов и алгоритмов, и полученных результатов.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИОННОЙ РАБОТЫ

Основные научно-теоретические и практические результаты, полученные при решении задач в рамках диссертационной работы, заключаются в следующем:

1. Проанализированы существующие методы, выявлены актуальные проблемы для обеспечения киберустойчивости КФС: разработка методов и алгоритмов интеллектуальной обработки больших данных, разработка методов и алгоритмов обеспечения кибербезопасности ИТ и защищенности ОТ КФС, разработка метода определения критичности уязвимостей функциональных компонентов КФС и метода оценки рисков КФС. Представлена постановка научной задачи исследования. Разработана трехуровневая концептуальная модель для решения ряда задач по обеспечению киберустойчивости КФС.

2. Разработаны алгоритмы, основанные на кластеризации с учетом компактности и разделимости кластеров, для обнаружения выбросов в больших данных на основе метода k-средних с целью снижения риска блокировки КФС.

3. Предложен алгоритм анализа больших данных КФС на основе взвешенной кластеризации, где веса, полученные путем суммирования весов каждой точки из набора данных, были назначены кластерам, что точнее обнаруживает аномалии в КФС по сравнению с алгоритмом k-средних.

4. Разработан метод параллельной обработки больших данных для снижения риска выхода КФС из строя, который сократил время кластеризации по сравнению с алгоритмом k-средних.

5. Предложен метод кластерного анализа больших данных КФС на основе консенсусного ансамбля с использованием квадрата евклидовой метрики.

6. Разработан метод классификации кибератак на КФС с применением экстремального машинного обучения и генетического алгоритма поведения «светлячков».

7. Разработан метод обнаружения DoS атак на КФС с применением ансамбля из пяти классификаторов, где полученная

метка класса соответствует максимальному значению среди всех оценок, полученных методами классификации для данной точки [10, 12].

8. Предложен алгоритм обнаружения вредоносного ПО в КФС на основе изображений с применением глубоких нейронных сетей.

9. Разработан метод обнаружения аномалий в ОТ КФС на основе иерархических скрытых Марковских моделей для системы водоочистки.

10. Разработан метод обнаружения аномалий в акустических сигналах КФС с применением трансферного обучения и метода экстремального градиентного бустинга [26].

11. Предложен метод обнаружения атак на устройства КФС с применением глубокой гибридной модели для системы водоочистки.

12. Разработан метод классификации отказов КФС на основе изображений с применением глубоких нейронных сетей на примере нефтяного погружного насоса.

13. Разработан метод определения критичности уязвимостей функциональных компонентов КФС на основе Байесовского графа атак для транспортировки природного газа по трубопроводу.

14. Разработан метод оценки рисков КФС с применением нечеткого интеграла Сугено на примере системы генерации энергии ветра.

Результаты, полученные при экспериментальных исследованиях, показали преимущества разработанных методов и доказывают перспективы выбранных подходов.

Основные результаты диссертации опубликованы в следующих работах:

1. Imamverdiyev, Y.N., Sukhostat, L.V. Anomaly detection in network traffic using extreme learning machine // IEEE International Conference on Application of Information and Communication Technologies (AICT'2016). Baku, Azerbaijan, – 2016, – p. 418-421. (**WoS, Scopus**)
2. Имамвердиев, Я.Н., Сухостат, Л.В. Вопросы безопасности киберфизических систем // "Riyaziyyatın tətbiqi məsələləri və yeni informasiya texnologiyaları" üzrə III respublika elmi-praktiki konfransı. Сумгаит, Азербайджан, – 2016, – с. 257-259.
3. Имамвердиев, Я.Н., Сухостат, Л.В. Вопросы применения методов машинного обучения для решения проблем информационной безопасности // "Big data: imkanları, multidissiplinar problemləri və perspektivləri" I respublika elmi-praktiki konfransı. Баку, Азербайджан, – 2016, – с. 127-131.
4. Алгулиев, Р.М. Киберфизические системы: основные понятия и вопросы обеспечения безопасности / Р.М.Алгулиев, Я.Н.Имамвердиев, Л.В.Сухостат // Информационные технологии, – 2017, 7, – с. 517–528. (**РИНЦ**)
5. Alguliyev, R.M. Anomaly Detection in Big Data based on Clustering / R.Alguliyev, R.Aliguliyev, Y.Imamverdiyev [et al.] // Statistics, Optimization and Information Computing, – 2017, 5 (4), – p. 325-340. (**Scopus (Q3)**)
6. Имамвердиев, Я.Н. Обнаружение аномалий в сетевом трафике на основе информативных признаков / Я.Н.Имамвердиев, Л.В.Сухостат // Radio Electronics, Computer Science, Control, – 2017, 3, – с. 113-120. (**WoS (Q4)**)
7. Alguliyev, R. An anomaly detection based on optimization / R.Alguliyev, R.Aliguliyev, Y.Imamverdiyev [et al.] // International Journal of Intelligent Systems and Applications, – 2017, 9 (12), – p. 87-96. (**Scopus (Q4))**
8. Алыгулиев, Р.М., Имамвердиев, Я.Н., Сухостат, Л.В. Оптимизационный подход к обнаружению аномалий в Big data // XIII Международная научно-техническая конференция «Распознавание-2017». Курск, Россия, – 2017, – с. 38-40. (**РИНЦ**)
9. Алгулиев, Р.М., Имамвердиев, Я.Н., Сухостат, Л.В. Обеспечение информационной безопасности киберфизических систем // "Proqram mühəndisliyinin aktual elmi-praktiki problemləri" I respublika konfransı. Баку, Азербайджан, – 2017, – с. 40-45.
10. Алгулиев Р.М., Алыгулиев Р.М., Имамвердиев Я.Н., Сухостат Л.В. Обнаружение DoS атак с применением ансамбля классификаторов // "İnformasiya təhlükəsizliyinin multidissiplinar problemləri" III respublika elmi-praktiki konfransı. Баку, Азербайджан, – 2017, – с. 12-18.
11. Alguliyev, R. Cyber-physical systems and their security issues / R.Alguliyev, Y.Imamverdiyev, L.Sukhostat // Computers in Industry, – 2018, 100, – p. 212-223. (**WoS (IF=4.769, Q1), Scopus (Q1)**)

12. Alguliyev, R.M. An improved ensemble approach for DoS attacks detection / R.M.Alguliyev, R.M.Aliguliyev, Y.N.Imamverdiyev [et al.] // Radio Electronics, Computer Science, Control, – 2018, 2 (45), – p. 73-82. (**WoS (Q4)**)
13. Alguliyev, R. Weighted clustering for anomaly detection in big data / R.Alguliyev, R.Aliguliyev, Y.Imamverdiyev [et al.] // Statistics, Optimization & Information Computing, – 2018, 6 (2), – p. 178–188. (**Scopus (Q3)**)
14. Alguliyev, R.M., Aliguliyev, R.M., Sukhostat, L.V. Purity-based consensus clustering for anomaly detection in Big data // XIV Internatıonal scientific conference "Recognition - 2018". Kursk, Russia, – 2018, – p. 17-20. (**РИНЦ**)
15. Сухостат, Л.В. Обнаружение атак на киберфизические системы на основе глубокого обучения // "İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri" IV respublika konfransı. Баку, Азербайджан, – 2018, – с. 42-46.
16. Alguliyev, R.M., Aliguliyev, R.M., Sukhostat, L.V. Consensus clustering by weight optimization of input partitions // IEEE 13th International Conference "Application of Information and Communication Technologies (AICT'2019). Baku, Azerbaijan, – 2019, – p. 1-4. (**WoS, Scopus**)
17. Алгулиев, Р.М., Имамвердиев, Я.Н., Шыхалиев, Р.Г., Сухостат, Л.В. Об одном подходе по выявлению кибератак на киберфизические системы с применением глубокой гибридной модели // The First International Scientific and Practical Forum «GLOBAL CYBER SECURITY FORUM 2019». Харьков, Украина, – 2019, – с. 27-28.
18. Alguliyev, R.M., Imamverdiyev, Y.N., Sukhostat, L.V. Detection of cyber-attacks on cyber-physical systems using deep neural network. XV International scientific conference "Recognition - 2019". Kursk, Russia, – 2019, – p. 18-20. (**РИНЦ**)
19. Сухостат, Л.В. Вопросы защиты персональных данных в киберфизических системах // "İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri" V respublika konfransı. Баку, Азербайджан, – 2019, – с. 92-96.
20. Alguliyev, R.M. Intelligent diagnosis of petroleum equipment faults using a deep hybrid model / R.M.Alguliyev, Y.N.Imamverdiyev, L.V.Sukhostat // SN Applied Sciences, – 2020, 2 (5), – p. 1-16. (**WoS (Q2), Scopus (Q3)**)
21. Alguliyev, R.M. Efficient algorithm for big data clustering on single machine / R.M.Alguliyev, R.M.Aliguliyev, L.V.Sukhostat // CAAI Transactions on Intelligence Technology, – 2020, 5 (1), – p. 9-14. (**WoS (Q2), Scopus (Q1)**)
22. Alguliyev, R.M. Weighted Consensus Clustering and its Application to Big Data / R.M.Alguliyev, R.M.Aliguliyev, L.V.Sukhostat // Expert Systems with Applications, – 2020, 150, – p. 1-15. (**WoS (IF=6.954, Q1), Scopus (Q1)**)
23. Alguliyev, R.M., Imamverdiyev, Y.N., Sukhostat, L.V. Diagnostics of DoS attacks on cyber-physical systems based on hierarchical hidden Markov models // XIII International Conference for young researchers "TECHNICAL SCIENCES. INDUSTRIAL MANAGEMENT". Borovets, Bulgaria, – 2020, – p. 39-41.

24. Alguliyev, R.M. Parallel batch k-means for big data clustering / R.M.Alguliyev, R.M.Aliguliyev, L.V.Sukhostat // Computers and Industrial Engineering, – 2021, 152, – p. 1-11. (**WoS (IF=7.180, Q1), Scopus (Q1)**)
25. Alguliyev, R.M. Hybrid DeepGCL model for cyber-attacks detection on cyber-physical systems / R.M.Alguliyev, Y.N.Imamverdiyev, L.V.Sukhostat // Neural Computing and Applications, – 2021, 33 (16), – p. 10211-10226. (**WoS (IF=5.102, Q2, Scopus (Q1)**)
26. Sukhostat, L.V. An intelligent model based on deep transfer learning for detecting anomalies in cyber-physical systems / L.V.Sukhostat // Radio Electronics, Computer Science, Control, – 2021, 3, – p. 124-132. (**WoS (Q4)**)
27. Сухостат, Л.В. Об одном подходе по обнаружения аномалий в киберфизических системах на основе акустических сигналов // 1st International Conference on "Information security: problems and prospects". Баку, Азербайджан, – 2021, – с. 115-118.
28. Сухостат, Л.В. Обзор некоторых решений безопасности современных АСУ ТП / Л.В.Сухостат // Телекоммуникации, – 2022, 2, – с. 14-23. (**РИНЦ**)
29. Sukhostat, L.V. Anomaly detection in industrial control system based on the hierarchical hidden Markov model / Cybersecurity for Critical Infrastructure Protection via Reflection of Industrial Control Systems, NATO Science for Peace and Security Series D: Information and Communication Security (IOS Press), – 2022, 62, – p. 48-55.
30. Алгулиев, Р.М., Алыгулиев, Р.М., Сухостат, Л.В. Оценка критичности киберфизических систем на основе графа атак // XVII International scientific conference "Recognition - 2023". Курск, Россия, – 2023, – с. 52-54. (**РИНЦ**)
31. Alguliyev, R. Radon transform based malware classification in cyber-physical system using deep learning / R.Alguliyev, R.Aliguliyev, L.Sukhostat // Results in Control and Optimization, – 2024, 14, – p. 1-14. (**WoS (IF=3.2, Q1), Scopus (Q1)**)
32. Alguliyev, R.M. Method for Quantitative Risk Assessment of Cyber-Physical Systems based on Vulnerability Analysis / R.M.Alguliyev, R.M.Aliguliyev, L.V.Sukhostat // Kybernetika, – 2024, 60 (6), – p. 779-796. (**WoS (IF=2.2, Q3), Scopus (Q3)**)
33. Alguliyev, R. An approach for assessing the functional vulnerabilities criticality of CPS components / R.Alguliyev, R.Aliguliyev, L.Sukhostat // Cyber Security and Applications, – 2025, 3, – p. 1-7. (**Scopus (Q1)**)

Роль соискателя в трудах, опубликованных в соавторстве:

[2-4,9,11] – проведен анализ методов и алгоритмов обеспечения киберустойчивости и информационной безопасности КФС;

[21,24] – предложены методы параллельной обработки больших данных для снижения риска выхода КФС из строя;

[13] – предложен алгоритм анализа больших данных КФС на основе взвешенной кластеризации;

[5,7,8] – предложены алгоритмы обнаружения выбросов в больших данных на основе метода k-средних для снижения риска блокировки КФС;

[14,16,22] – предложен метод кластерного анализа больших данных КФС на основе консенсусного ансамбля;

[1,6] – предложен метод классификации кибератак на КФС с применением экстремального машинного обучения;

[10,12] – предложен метод обнаружения DoS атак на КФС с применением ансамбля классификаторов;

[31] – предложен алгоритм обнаружения вредоносного ПО в КФС на основе изображений;

[23] – предложен метод обнаружения аномалий в ОТ КФС на основе иерархических скрытых Марковских моделей;

[17,18,25] – предложен метод обнаружения атак на устройства КФС с применением глубокой гибридной модели;

[20] – предложен метод классификации отказов КФС на основе изображений;

[30,33] – предложен метод определения критичности уязвимостей функциональных компонентов КФС на основе Байесовского графа атак;

[32] – предложен метод оценки рисков КФС с применением нечеткого интеграла Сугено.

Защита диссертации состоится 31 октября 2025 года в 10^{00} на заседании Диссертационного Совета ED 1.35 действующего на базе Института Информационных Технологий Министерства Науки и Образования Азербайджанской Республики.

Адрес: AZ1141, г.Баку, ул. Б.Вахабзаде, 9А.

С диссертацией можно ознакомиться в библиотеке Института Информационных Технологий Министерства Науки и Образования Азербайджанской Республики.

Электронные версии диссертации и автореферата размещены на официальном сайте (*ict.az*) Института Информационных Технологий Министерства Науки и Образования Азербайджанской Республики.

Автореферат разослан по соответствующим адресам 29 сентября 2025 года.

Подписано в печать: 27.09.2025

Формат бумаги: 60x84 1/16

Объем: 76033 (знаков)

Тираж: 70