Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Уфимский университет науки и технологий»

На правах рукописи

ШАМСУТДИНОВ РИНАТ РУСТЕМОВИЧ

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА МОНИТОРИНГА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ПРОМЫШЛЕННОГО ИНТЕРНЕТА ВЕЩЕЙ С ИСПОЛЬЗОВАНИЕМ МЕХАНИЗМОВ ИСКУССТВЕННЫХ ИММУННЫХ СИСТЕМ

Специальность 2.3.6. Методы и системы защиты информации,
информационная безопасность

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор технических наук, профессор
Васильев Владимир Иванович

УФА – 2023

Оглавление

Введение

Актуальность темы исследования. Современный этап развития промышленного производства (Индустрия 4.0) связан с массовым внедрением информационных технологий и киберфизических систем, масштабной автоматизацией бизнес-процессов, распространением технологий искусственного интеллекта. Ключевые позиции в нарождающейся 4-й промышленной революции начинает занимать промышленный Интернет вещей (Industrial Internet of Things, IIoT). Промышленный Интернет вещей представляет собой систему объединенных компьютерных сетей и подключенных к ним промышленных (производственных) объектов со встроенными датчиками и программным обеспечением (ПО) для сбора и обмена данными, с возможностью удаленного контроля и управления в автоматизированном режиме, без участия человека. По прогнозам, к 2030 г. этот сектор будет добавлять к мировому валовому продукту (ВВП) 14,2 трлн. долл., а количество подсоединенных к сети устройств составит 75.4 млрд.

Важным условием работы сетей и систем промышленного Интернета вещей является обеспечение их информационной безопасности (ИБ). Согласно данным компании Check Point, число кибератак за 2022 г. выросло на 38%. 94% производственных компаний столкнулись с инцидентами ИБ в 2022 г., при этом менее трети из них внедрили на своих объектах системы безопасности IIoT и операционных технологий (Operational technologies, OT). Наиболее распространенные причины слабой защищенности IIoT: устаревшее ПО и недостаточное внимание к программным обновлениям, передача данных без шифрования, стандартные заводские настройки безопасности устройств IIoT, незащищенные интерфейсы, уязвимости в операционных системах общего назначения, невозможность оснастить многие устройства встроенными средствами безопасности и т.п. В качестве средств мониторинга ИБ сетей используются системы обнаружения атак и аномалий, применяемые в том числе в промышленных автоматизированных системах (АСУ ТП), но существующие

системы обнаружения сетевых атак и аномалий, не в полной мере учитывают специфику IIoT.

Невысокая эффективность существующих систем обнаружения атак (СОА) в решении задачи мониторинга ИБ сетей IIoT обуславливается следующим. Существуют атаки, эксплуатирующие специфичные уязвимости IIoT, к примеру, атаки, нацеленные на повышение расхода электроэнергии устройствами IIoT, в том числе сенсоров беспроводных сенсорных сетей, не выявляемые штатными средствами защиты промышленных систем. Существующие СОА, основанные на анализе сигнатур, в принципе не могут выявлять новые, не известные ранее атаки (атаки нулевого дня), а также новые виды уже известных атак, для которых пока не существует сигнатур. СОА на основе выявления злоупотреблений (аномалий) характеризуются большим количеством допускаемых ошибок. Существующие СОА, реализующие в том числе методы и алгоритмы искусственного интеллекта, демонстрируют более высокую, но всё еще недостаточную эффективность обнаружения сетевых атак и аномалий в сетях IIoT.

Поэтому проблема разработки и исследования новых методов и алгоритмов к обнаружению атак и аномалий в работе систем и устройств IIoT с целью повышения их уровня защищенности в условиях воздействия возможных внешних и внутренних угроз безопасности информации является актуальной.

Степень разработанности темы исследования. Исследованиям в области оценки рисков ИБ и уровня защищенности промышленных автоматизированных систем (АСУ ТП) посвящены работы таких российских и зарубежных ученых, как: Ажмухамедов И.М., Аникин И.В., Васильев В.И., Вульфин А.М., Катасёв А.С., Костогрызов А.И., Котенко И.В., Машкина И.В., Шелупанов А.А., Massacci F., Noel S. и др. Конкретные технические решения в области построения систем обнаружения сетевых атак и аномалий предложены в работах таких авторов, как: Абрамов Е.С., Аралбаев Т.З., Бурлаков М.Е., Карташевский В.Г., Поздняк И.С., Aydin M., Dilek S., Çakır H., He J., Jing C., Li Y., Powers S.T., Xu J. и др. Достаточно большое количество работ посвящено решению задач, связанных с реализацией процессов мониторинга ИБ в IIoT-системах, в том числе с

применением методов и технологий искусственного интеллекта. Исследования в данном направлении отражены в публикациях таких авторов, как: Заводцев И.В., Зегжда Д.П., Глухих И.Н., Дойникова Е.В., Милославская Н.Г., Полтавцева М.А., Сычугов А.А., Тебуева Ф.Б., Alaparthy V., Gupta G.P., Joshi A., Morgera S., Raja K., Singh A., Xiao X., Zhang R., Zou Z. и др.

Вместе с тем, существующие методики обнаружения атак и аномалий сетевого трафика IIoT-систем нуждаются в развитии и совершенствовании, что может быть реализовано на основе применения технологий искусственного интеллекта, включая применение искусственных иммунных систем, чему и посвящена настоящая работа.

Объект исследования – сети и системы промышленного Интернета вещей.

Предмет исследования – защищённость сетей промышленного Интернета вещей от угроз информационной безопасности.

Методы исследования. В работе использовались методы системного анализа, защиты информации, машинного обучения, искусственных иммунных систем, нейронных сетей, многоагентных систем, SIEM-систем, компьютерного моделирования.

Цель работы – повышение эффективности систем мониторинга ИБ сетей промышленного Интернета вещей на основе разработки и применения алгоритмов обнаружения компьютерных атак и аномалий сетевого трафика с использованием механизмов искусственных иммунных систем.

Задачи исследования:

1. Анализ современного состояния исследований в области мониторинга ИБ сетей промышленного Интернета вещей с использованием технологий искусственного интеллекта.
2. Разработка алгоритмов обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием адаптивных механизмов искусственных иммунных систем.

3. Разработка алгоритмов обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием ансамбля методов машинного обучения и подсистемы корреляции событий ИБ.

4. Разработка архитектуры исследовательского прототипа интеллектуальной системы мониторинга ИБ промышленного Интернета вещей, оценка эффективности ее применения при решении прикладных задач.

Положения, выносимые на защиту.

1. Результаты анализа современного состояния исследований в области мониторинга сетей промышленного Интернета вещей с использованием технологий искусственного интеллекта, предложенная концепция построения интеллектуальной системы мониторинга ИБ промышленного Интернета вещей.

2. Алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием адаптивных механизмов искусственных иммунных систем.

3. Алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием ансамбля методов машинного обучения и подсистемы корреляции событий ИБ.

4. Архитектура исследовательского прототипа интеллектуальной системы мониторинга ИБ промышленного Интернета вещей, результаты ее применения при решении прикладных задач.

Научная новизна полученных результатов

1. Предложена концепция построения интеллектуальной системы мониторинга информационной безопасности (ИБ) промышленного Интернета вещей на основе многоагентной платформы гибридной многоуровневой интеллектуальной системы обнаружения атак и аномалий сетевого трафика IIoT, отличающаяся интеграцией механизмов искусственных иммунных систем, методов машинного обучения, подсистемы корреляции событий информационной безопасности (SIEM-системы), что позволяет повысить полноту и точность выявления внешних и внутренних угроз безопасности информации.

2. Разработаны алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей на основе адаптивных механизмов искусственных иммунных систем (ИИС), отличающиеся интеграцией различных подходов в рамках теории ИИС (негативная селекция, клональный отбор, теория опасности, теория иммунной сети) и модификации известных алгоритмов клональной селекции, дендритных клеток и обновления детекторов, что позволяет в совокупности существенно снизить уровень принятия ошибочных решений, выявлять в том числе новые, ранее неизвестные системе угрозы безопасности информации.

3. Разработаны алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей на основе использования ансамбля методов машинного обучения и подсистемы корреляции событий ИБ, отличающиеся использованием специфического иммунного ответа ИИС совместно с другими алгоритмами искусственного интеллекта (нейронные сети, алгоритм случайного леса), что обеспечивает дифференцированный подход к обнаружению различных типов атак и аномалий (включая ранее неизвестные) и позволяет повысить эффективность функционирования системы мониторинга ИБ в целом.

4. Разработана архитектура исследовательского прототипа интеллектуальной системы мониторинга ИБ промышленного Интернета вещей на основе комитета классификаторов, отличающаяся использованием многоагентного подхода к построению многоуровневых распределенных гибридных интеллектуальных систем, что позволяет более полно использовать преимущества применения различных технологий интеллектуального анализа данных, учесть особенности структурно-функциональной организации (состава подсистем) объекта мониторинга, многообразие угроз безопасности информации и уязвимостей программного обеспечения, дополняя полученную информацию текущими данными от взаимодействующей SIEM-системы.

Теоретическая значимость полученных результатов заключается в том, что в работе предложены: концепция построения интеллектуальной системы

мониторинга ИБ промышленного Интернета вещей, алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием адаптивных механизмов ИИС; алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием ансамбля методов машинного обучения и подсистемы корреляции событий ИБ; архитектура исследовательского прототипа интеллектуальной системы мониторинга ИБ промышленного Интернета вещей.

Практическая значимость полученных результатов заключается в разработке: программных модулей исследовательского прототипа интеллектуальной системы мониторинга ИБ промышленного Интернета вещей; методики ее применения для обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей. Применение предложенных модулей позволяет обеспечить точность обнаружения компьютерных атак и аномалий сетевого трафика на уровне 98-99% на тестовых наборах данных, выявлять в том числе новые, ранее неизвестные системе угрозы безопасности информации. Взаимодействия предложенных решений с существующими SIEM-системами позволяет при этом дополнительно повысить уровень достоверности принимаемых решений в процессе мониторинга ИБ сетей промышленного Интернета вещей.

Достоверность результатов работы обеспечивается корректным использованием основных теоретических положений, методов проведения вычислительных экспериментов, непротиворечивостью полученных результатов, а также их экспертной оценкой и степенью повторяемости, апробацией на научных конференция, публикацией результатов в ведущих рецензируемых научных изданиях из Перечня ВАК, а также в научных изданиях, индексируемых в Scopus и Web of Science.

Связь исследований с научными программами. Исследования выполнялись в рамках грантов РФФИ № 20-37-90024 «Гибридная интеллектуальная система мониторинга информационной безопасности на основе алгоритмов искусственных иммунных систем и нечетких нейронных сетей»,

№ 20-08-00668 «Разработка и исследование методологии, моделей и методов комплексного анализа и управления рисками кибербезопасности АСУ ТП промышленных объектов с использованием технологии когнитивного моделирования и интеллектуального анализа данных».

Соответствие паспорту специальности. Результаты диссертационной работы соответствуют следующим пунктам паспорта научной специальности 2.3.6 «Методы и системы защиты информации, информационная безопасность»: п. 3 «Методы, модели и средства выявления, идентификации, классификации и анализа угроз нарушения информационной безопасности объектов различного вида и класса» – разработаны методика и алгоритмы обнаружения и классификации компьютерных атак, представляющих угрозы безопасности информации сетей промышленного Интернета вещей; п. 6 «Методы, модели и средства мониторинга, предупреждения, обнаружения и противодействия нарушениям и компьютерным атакам в компьютерных сетях» – разработана архитектура исследовательского прототипа интеллектуальной системы мониторинга ИБ сетей промышленного Интернета вещей; п. 15 «Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности» – разработаны алгоритмы обнаружения компьютерных атак и сетевых аномалий, архитектура исследовательского прототипа интеллектуальной системы мониторинга для усовершенствования существующих систем обнаружения вторжений, систем мониторинга и применения в качестве новых, отдельных решений.

Апробация результатов.

Полученные результаты исследований докладывались и обсуждались на следующих научных конференциях: VII Всероссийская заочная Интернет-конференция «Проблемы информационной безопасности», г. Ростов-на-Дону, 2018 г.; VI – IX Всероссийские научные конференции (с приглашением зарубежных ученых) «Информационные технологии интеллектуальной поддержки принятия решений », гг. Уфа, Ставрополь, Ханты-Мансийск, 2018-

2021 гг.; XIII Всероссийская молодёжная научная конференция «Мавлютовские чтения», г. Уфа, 2019 г.; 13-я Всероссийская зимняя школа-семинар магистрантов, аспирантов и молодых ученых «Актуальные проблемы науки и техники», г. Уфа, 2020 г.; XIV Всероссийская молодежная научная конференция «Мавлютовские чтения», г. Уфа, 2020 г.; Международная научная конференция «Цифровая индустрия: состояние и перспективы развития 2020», г. Челябинск, 2020 г.; Студенческая научно-техническая конференция «Неделя науки», г. Уфа, 2021 г.; Всероссийский конкурс-конференция студентов и аспирантов по информационной безопасности «SIBINFO-2021», г. Томск, 2021 г.

Результаты диссертационной работы внедрены в ЗАО «Республиканский центр защиты информации» и в учебный процесс кафедры вычислительной техники и защиты информации ФГБОУ ВО «Уфимский университет науки и технологий» Министерства науки и высшего образования РФ.

Публикации по теме исследования

По проблеме диссертационного исследования опубликовано 17 работ, в том числе: 6 в рецензируемых научных журналах, входящих в перечень изданий, рекомендованных ВАК, 3 – в изданиях, индексируемых в Web of Science, одно из них индексируется в Scopus; 8 – в прочих изданиях.

Структура и объем работы. Работа состоит из введения, четырех глав, заключения, списка сокращений и условных обозначений, словаря терминов, списка литературы и приложений, содержит 38 рисунков, 38 таблиц, 26 формул, 3 приложения, выполнена на 187 страницах. Список использованной литературы включает 201 источник.

1 Анализ современного состояния исследований в области автоматизации мониторинга информационной безопасности сетей промышленного Интернета вещей с использованием технологий искусственного интеллекта

1.1 Проблема обеспечения информационной безопасности сетей промышленного Интернета вещей

Современный этап развития промышленности (Industry 4.0) в значительной степени связан с развитием и внедрением промышленного Интернета вещей (Industrial Internet of Things, IIoT). По оценкам аналитической компании ResearchandMarkets, ежегодные темпы роста мирового рынка промышленного Интернета вещей, начиная с 2021 г., составляют в среднем 21%, а объем мирового рынка IIoT к 2026 г. должен достигнуть 344,7 млрд. долл.

Промышленный Интернет вещей является расширением Интернета вещей (Internet of Things, IoT), направленным на выполнение промышленных задач. Под Интернетом вещей (IoT) обычно понимается концепция вычислительной сети физических устройств («вещей»), оснащенных встроенными технологиями для взаимодействия друг с другом или с внешней средой [139]. С точки зрения технологий, Интернет вещей представляет собой 4-хзвенную систему: подключаемые устройства (сенсоры, датчики, терминалы); сети, по которым они взаимодействуют; IoT-платформы и приложения для конечных пользователей.

Согласно предварительному национальному стандарту ПНСТ 643-2022. «Информационные технологии. Интернет вещей промышленный. Термины и определения» [10], промышленный Интернет вещей (IIoT) – это Интернет вещей, машин, компьютеров и людей, обеспечивающий интеллектуальные производственные операции с использованием расширенной аналитики данных для качественно новых результатов бизнеса. В промышленном Интернете вещей основными разновидностями «вещей», которые надо подключать к сети, являются различные типы датчиков (сенсоров) и приводов. Эти устройства, с одной

стороны, имеют интерфейс с коммуникационной сетью, а с другой – интерфейс, обеспечивающий физическое взаимодействие с процессом, который требуется отслеживать. Коммуникационный интерфейс является необходимой компонентой IIoT. Это может быть проводной или беспроводной интерфейс. Но, независимо от того, какая технология используется на канальном и физическом уровнях, устройства должны поддерживать протокол IP, чтобы интегрироваться в инфраструктуру IIoT.

Понятие IIoT тесно связано с понятиями киберфизической системы и АСУ ТП. Предварительный национальный стандарт ПНСТ 417-2020 «Система киберфизическая. Термины и определения» [8] определяет киберфизическую систему как «интеллектуальную систему, которая включает в себя инженерные взаимодействующие сети физических и вычислительных компонентов». Под АСУ ТП традиционно понимается целостное решение, обеспечивающее автоматизацию основных операций ТП на производстве в целом или каком-то его участке, выпускающем относительно завершенное изделие, хотя в последнее время функции и целевое назначение АСУ ТП значительно сместились в сторону идеологии IIoT. Отметим также, что многие промышленные объекты в нашей стране, использующие или внедряющие технологии IIoT, подпадают под действие федерального закона РФ № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации» [1]. В соответствии с этим законом, к объектам критической информационной инфраструктуры (КИИ) относятся информационные системы, информационно-телекоммуникационные сети, автоматизированные системы управления и другие значимые системы, функционирование которых критически важно для жизнедеятельности государства.

В 2020 году в РФ был разработан предварительный национальный стандарт ПНСТ 420-2020 «Информационные технологии. Интернет вещей промышленный. Типовая архитектура» [9], который определяет типовую архитектуру промышленного Интернета вещей, структуру архитектуры IIoT, включающую каркас, основанный на интересах, точках зрения, видах моделей,

заинтересованных сторонах. Схема типовой архитектуры IIoT и ее применение, предлагаемые стандартом, представлены на Рисунке 1.1.

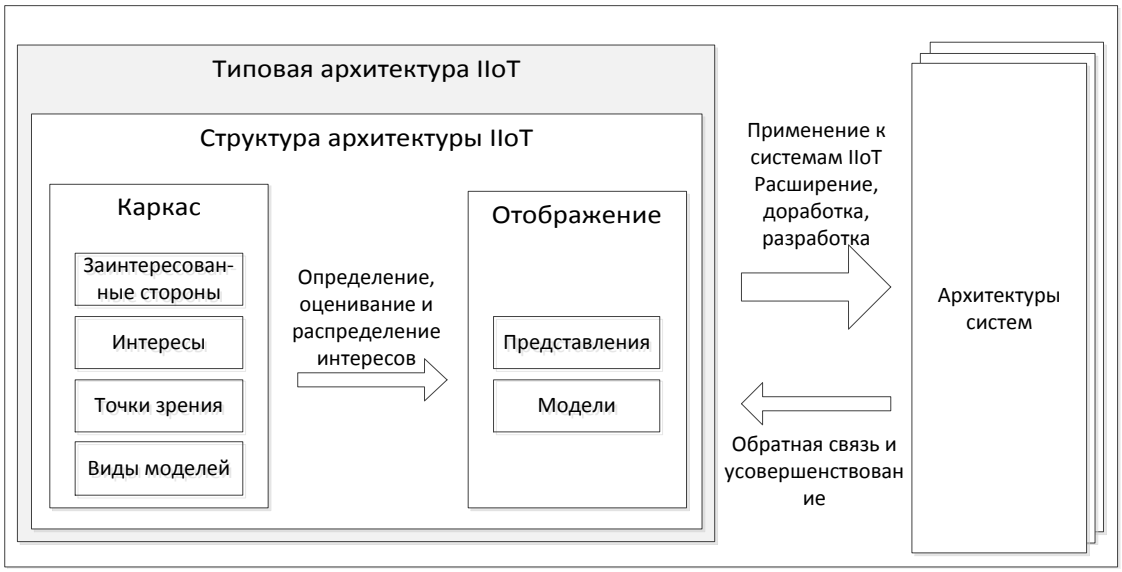


Рисунок 1.1 – Схема типовой архитектуры IIoT и ее применение [9]

Пример трехуровневой архитектуры промышленного Интернета вещей, представленной в стандарте с точки зрения реализации архитектурных паттернов (шаблонов), приведен на Рисунке 1.2. Данная архитектура включает в себя следующие уровни:

– уровень предприятия, где реализованы приложения, системы поддержки принятия решения, интерфейсы для конечных пользователей (здесь осуществляется получение потоков информации с других уровней и выдача управляющих команд);

– уровень платформы, где агрегируется и обрабатывается информация граничного уровня, перенаправляются команды управления с уровня предприятия на граничный уровень;

– граничный уровень, где осуществляется сбор данных от граничных узлов посредством сети ближнего действия, а также реализация управляющих команд.

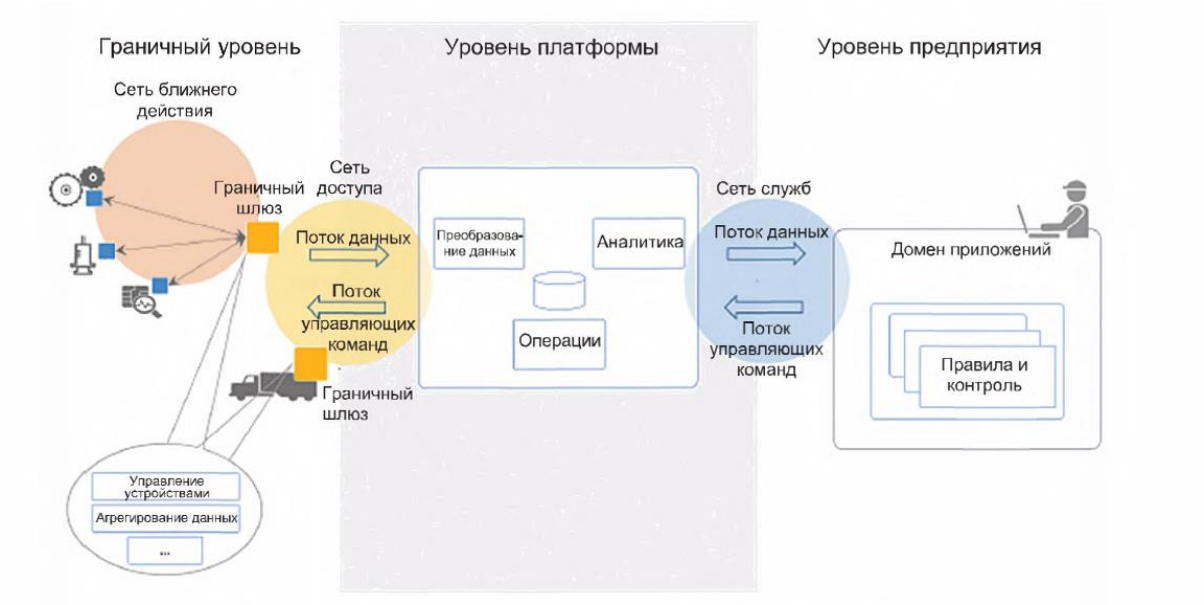


Рисунок 1.2 – Трёхуровневая архитектура реализации IIoT

Стандарт также выделяет функциональные области IIoT, представленные Рисунком 1.3.

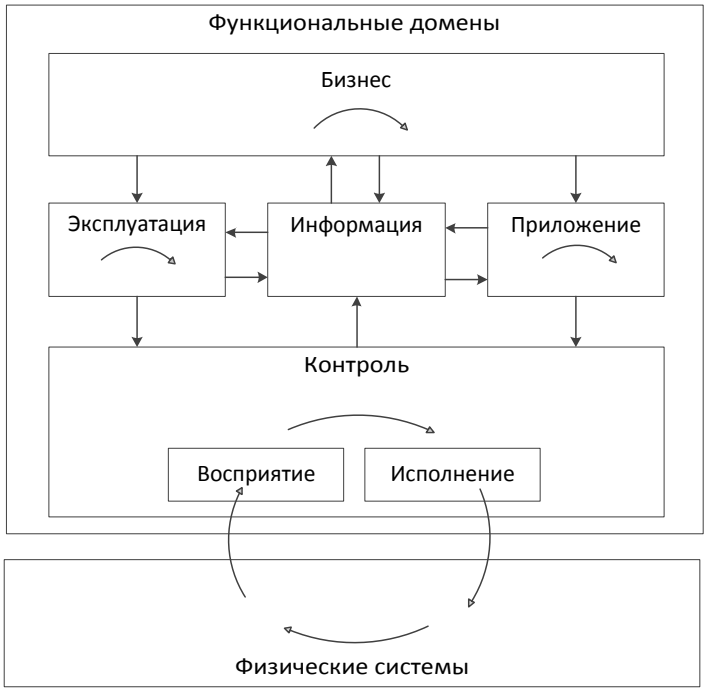


Рисунок 1.3 – Функциональные области IIoT

Представление функциональных доменов на основе трёхуровневой архитектуры IIoT показано на Рисунке 1.4.

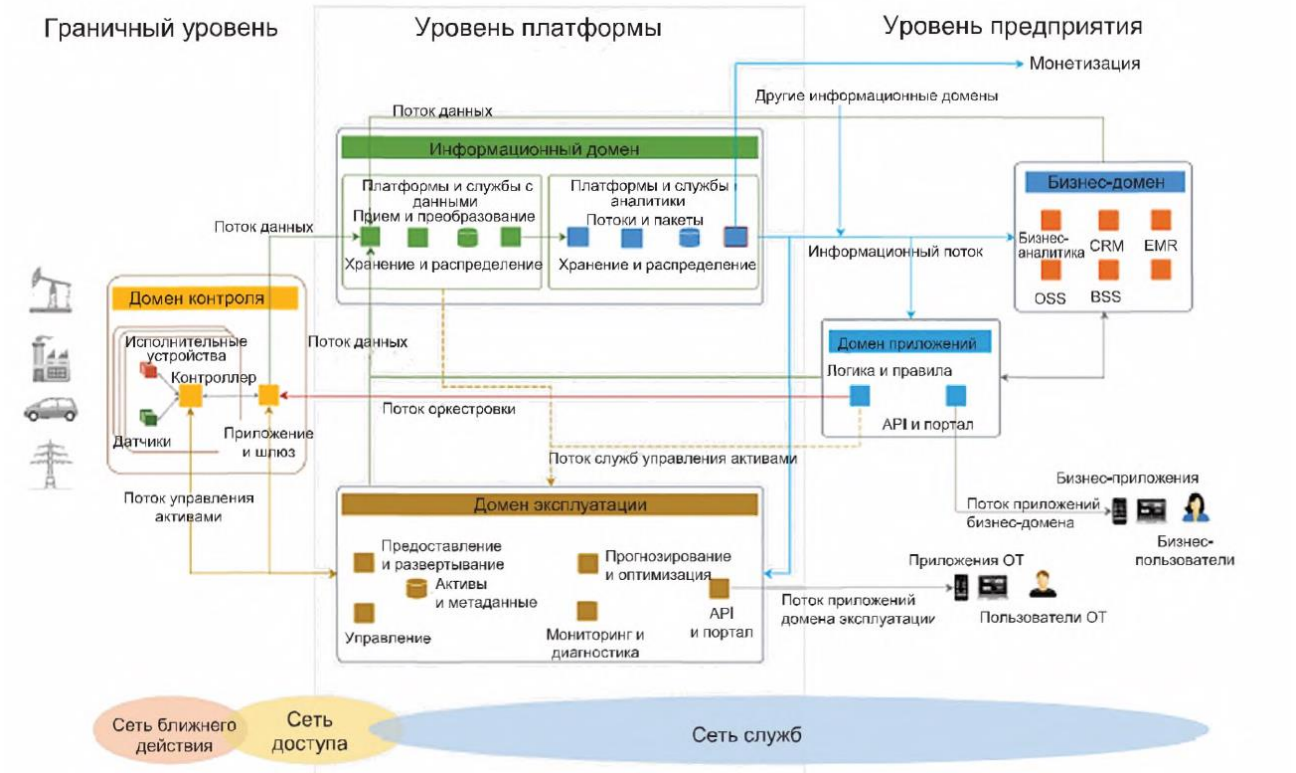


Рисунок 1.4 – Функциональные домены в трехуровневой архитектуре IIoT

Одной из главных проблем в построении и эксплуатации промышленного Интернета вещей является обеспечение ИБ его устройств и систем. Наиболее распространенные причины слабой защищенности IIoT:

– устаревшее системное и прикладное ПО устройств IIoT, недостаточное внимание к программным обновлениям;

– передача данных без шифрования;

– стандартные заводские настройки безопасности устройств;

– незащищенные интерфейсы;

– уязвимости в операционных системах (ОС) общего назначения;

– невозможность оснастить многие устройства встроенными средствами безопасности.

Специфика IIoT заключается в подключении промышленных систем к сети Интернет, возможности реализации удаленного управления ими, в том числе с устройств, находящихся за пределами предприятия; использовании облачных

систем, в том числе арендованных; ограниченности вычислительных и энергетических ресурсов автономных IIoT-устройств; их слабой защищенности; отсутствии, зачастую, средств шифрования трафика. В данной работе не рассматриваются вопросы шифрования и защиты облачных систем.

Согласно отчету фирмы Nokia «Threat Intelligence Report 2020» [185], в последние годы доля атак на устройства IIoT в общем числе атак на мобильные устройства увеличилась и достигла значения 32,7 %. По данным Лаборатории Касперского, количество новых образцов вредоносного ПО для IIoT-устройств с каждым годом значительно выросло: если в 2015 году их число составляло 483, то в 2020 г. – уже 331 401 [55]. 55% респондентов опроса [56], проведенного данной лабораторией, характеризуют IIoT как один из главных факторов, влияющих на кибербезопасность АСУ ТП, но в то же время только 14% организаций используют средства обнаружения сетевых аномалий и 19% – системы мониторинга сети и трафика. По данным Check Point [120], 67% предприятий сегодня уже столкнулись с инцидентами безопасности, связанными с применением IIoT-устройств

В [74] также подчеркивается уязвимость устройств промышленного Интернета вещей, в качестве слабых мест при этом отмечается продолжающийся переход на IPv6, слабые аутентификация и стандартные учетные записи, трудности обновления программного обеспечения (ПО) и отсутствие поддержки производителя, открытое состояние неиспользуемых портов, применение текстовых протоколов, незащищенных мобильных технологий, облачной инфраструктуры, уязвимого ПО и человеческий фактор.

Производители, конечно, стараются решать проблемы безопасности IIoT. Так, по данным [46], в целях обеспечения ИБ IoT в облачной системе Microsoft Azure используется методика моделирования угроз STRIDE, которая рассматривает все уровни и компоненты IoT с точки зрения возникновения различных угроз, предлагаются меры защиты. В [111] проводится сравнение возможностей обеспечения ИБ таких IoT-фреймворков, как AWS IoT от компании Amazon, ARM Bed от компании ARM, Azure IoT от «Microsoft», HomeKit от

«Apple», Brillo/Weave «Google», SmartThings от «Samsung», Calvin от «Ericsson» и Kura от «Eclipse».

В [55] отмечается, что предотвратить угрозу выгоднее, чем компенсировать ущерб от последствий ее реализации. Сообщается, что Лабораторией Касперского разработан первый IT-продукт с кибериммунитетом – KISG 100. Данный продукт представляет собой шлюз данных для IoT.

Необходимо отметить, что в рамках IIoT часто используются беспроводные сенсорные сети (Wireless Sensor Networks, WSN), состоящие из большого числа автономных сенсорных узлов, собирающих различные данные и обменивающихся ими при помощи беспроводного соединения с более мощным узлом – базовой станцией. В связи с их распределенной открытой архитектурой и ограниченностью ресурсов сенсорных узлов, такие сети являются очень уязвимыми для атак [174]. Согласно [130, 155], злоумышленник может скомпрометировать сенсорный узел, прослушивать, искажать и имитировать сообщения, нарушить целостность данных и повысить расход ресурсов. Одним из наиболее распространённых и опасных видов атак, угрожающих WSN, являются атаки «отказ в обслуживании» (Denial of Service, DoS).

В настоящее время ведутся работы по повышению уровня защищенности WSN. В [166] проанализированы функциональные особенности WSN и наиболее распространенные типы атак. В качестве защиты от вредоносных или скомпрометированных узлов предлагается использовать адаптивное взаимодействие элементов системы, основанное на анализе поведения соседних узлов.

В [196] подчеркивается высокая уязвимость сенсорных узлов WSN, возможность злоумышленников нанести большой ущерб при успешной компрометации узла; определены уязвимости алгоритмов обмена аутентификационной информацией. Во избежание компрометации узла авторы предлагают дополнить существующий алгоритм новой схемой обмена аутентификационной информацией. Оценка BAN-логикой, а также проведённые

оценки производительности и защищённости показывают эффективность предложенной схемы.

В [14] рассматривается вопрос разработки системы обнаружения атак для WSN, проанализированы схемы обнаружения на основе правил, интеллектуального анализа данных, теории игр и статистики, предлагается реализация гибридной системы обнаружения аномалий.

Важную роль в формировании нормативной базы обеспечения ИБ промышленного Интернета вещей сыграло принятие серии стандартов ГОСТ Р МЭК «Сети коммуникационные промышленные. Защищенность (кибербезопасность) сети и системы» – ГОСТ Р 56205-2014 / IEC/TS 62443-1-1:2009 [3], ГОСТ Р МЭК 62443-2-1-2015 [6], ГОСТ Р МЭК 62443-3-3-2016 [7]. В этих стандартах предложена методика оценки защищенности (кибербезопасности) промышленных коммуникационных сетей, выполнение которой предусматривает: выделение в составе исследуемого объекта (сети) относительно независимых зон безопасности и связывающих их каналов передачи информации (трактов); построение модели физической архитектуры, зональной и трактовой модели сети; оценка угроз и рисков ИБ с использованием этих моделей; определение целевого и достигнутого уровня безопасности; а также, в случае необходимости, принятие дополнительных контрмер по защите информации (ЗИ).

Среди международных документов в области обеспечения ИБ IoT, IIoT и промышленных автоматизированных систем можно выделить отчет NISTIR [94], рекомендации и практики Министерства внутренней безопасности США [97], Агентства Европейского союза по сетям и информационной безопасности (European Network and Information Security Agency, ENISA) [88-90], а также международные стандарты [91-93, 95-96, 98-100].

Следует отметить также принятие международными организациями по стандартизации ИСО и МЭК стандарта российской разработки ISO/IEC 30162:2022 «Интернет вещей. Требования к совместимости устройств, сетей и систем промышленного Интернета вещей» (Internet of Things (IoT) – Compatibility

requirements and model for devices within Industrial IoT systems) [69]. Формирование единых требований к совместимости различных устройств IIoT способствует решению проблемы интеграции этих устройств, обусловленной применением различными производителями своих собственных протоколов и стандартов, ввиду отсутствия ранее единого международного стандарта. Это, в свою очередь, способствует решению задачи обеспечения ИБ сетевого взаимодействия IIoT-устройств.

1.2 Методы обеспечения информационной безопасности сетей промышленного Интернета вещей

Требования к безопасности сетей промышленного Интернета вещей определяются требованиями к обеспечению безопасности систем, в рамках которых они функционируют, и безопасности информации, которую они обрабатывают.

Как уже отмечалось, системы IIoT имеют много общего с промышленными системами автоматизации (АСУ ТП) и объектами КИИ, поэтому большинство методов по обеспечению их ИБ будут обусловлены требованиями по ИБ, предъявляемыми к системам и объектам данного вида. Это, прежде всего, Приказы ФСТЭК №31 [11], №235 [12] и №239 [13].

Приказ ФСТЭК от 14.03.2014 г. № 31 «Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды» устанавливает конкретные требования к системам защиты АСУ ТП, правила по определению класса защищенности АСУ и соответствующий состав необходимых мер защиты. Подходы к обеспечению ИБ АСУ ТП на основе положений Приказа №31, включая применение

интеллектуальных систем анализа состояний рассматриваемых объектов, подробно рассмотрены в [87].

Приказ ФСТЭК России от 21.12.2017 г. №235 «Об утверждении Требований к созданию систем безопасности значимых объектов критической информационной инфраструктуры Российской Федерации и обеспечению ее функционирования» устанавливает требования к применяемым программным и программно-аппаратным комплексам, а также требования к функционированию системы безопасности в части организации работ по обеспечению безопасности и к организационно-распорядительным документам по безопасности значимых объектов.

Конкретные требования и состав необходимых мер по обеспечению ИБ значимых объектов КИИ устанавливаются Приказом ФСТЭК России от 25.12.2017 г. №239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации». Этим документом устанавливаются цели и задачи обеспечения безопасности, объекты защиты, состав мер по организационно-технической ЗИ, требования к обеспечению безопасности в ходе создания, эксплуатации и вывода из эксплуатации значимых объектов КИИ и их программно-аппаратных комплексов.

Программа стандартизации в области промышленного Интернета вещей пока находится в стадии разработки и обсуждения серии предварительных национальных стандартов (ПНСТ), некоторые из которых упоминались в разделе 1.1. Международные стандарты уже сегодня регулируют широкий круг вопросов, относящихся к обеспечению безопасности IIoT-систем [91-93, 95-96, 98-100].

Непосредственное отношение к тематике промышленного Интернета вещей имеет ГОСТ Р 59547-2021 «Защита информации. Мониторинг информационной безопасности. Общие положения» [4]. Этот стандарт описывает решаемые в рамках мероприятий по мониторингу ИБ задачи, объекты и уровни мониторинга, реализуемые свойства мониторинга: многопараметричность, масштабируемость, адаптивность, полнота, доступность и достоверность. Стандарт устанавливает

общие требования к мониторингу ИБ, а именно к источникам данных, сбору, хранению, агрегированию и обработке данных, требования к представлению данных о результатах мониторинга и их защите, а также к порядку осуществления мониторинга ИБ при реализации мер ЗИ.

Понятие риска в контексте информационной безопасности зачастую рассматривается в качестве количественной меры ущерба от реализации угрозы с учетом вероятности ее реализации. Прогнозирование и оценка рисков ИБ рассматривается в [15, 50, 186], для роевых робототехнических систем – в [70], методы управления рисками ИБ корпоративных сетей – в [19], критических инфраструктур – в [51], в составе системы менеджмента информационной безопасности (СМИБ) КИИ – в [58]. В целом, СМИБ и методика мониторинга событий рассматриваются в [85], модель активного мониторинга для СМИБ КФС – в [72].

Отметим, что в настоящее время с целью повышения уровня автоматизации процессов, связанных с управлением инцидентами ИБ, повышения эффективности реагирования на киберугрозы, обеспечения комплексной защиты компьютерных сетей создаются и используются ситуационные центры управления ИБ (Security Operation Center, SOC), рассматриваемые в [64].

По данным [22], большую часть времени оператор SOC-центра работает с системами управления безопасностью и событиями безопасности (Security Information and Event Management, SIEM). SIEM-системы осуществляют сбор данных о событиях по всей сети с различных источников, сопоставляют события между собой, выявляют подозрительные совокупности событий, которые вне этих совокупностей могут выглядеть вполне легитимными. На основе корреляционного анализа осуществляется более глубокая обработка данных о событиях и лучшее выявление инцидентов ИБ.

Подсистемы корреляционного анализа являются неотъемлемой составляющей SIEM (Security Information and Event Management) – систем, осуществляющих управление информацией о событиях и инцидентах ИБ. Методы

SIEM эффективно используются в том числе для обнаружения сетевых атак на IIoT [135].

Источники информации для SIEM-систем включают:

– системы аутентификации и контроля доступа, предоставляющие информацию об успешных или неуспешных попытках получения доступа;

– DLP-системы, представляющие сведения о попытках инсайдерских утечек, нарушении прав доступа;

– IDS/IPS-системы, предоставляющие данные о сетевых атаках, изменениях конфигурации и доступа к устройствам;

– межсетевые экраны, предоставляющие сведения об атаках, вредоносном ПО, попытках нарушения правил доступа и пр.;

– антивирусные приложения, генерирующие события о работоспособности ПО, баз данных, изменении конфигураций и политик, вредоносном коде;

– журналы событий серверов и рабочих станций, используемые для контроля доступа, соблюдения политик информационной безопасности;

– сетевое активное оборудование, используемое для контроля доступа, учета сетевого трафика;

– сканеры уязвимостей, предоставляющие данные об инвентаризации активов, сервисов, ПО, уязвимостей, топологической структуры;

– системы инвентаризации и управления активами, поставляющие данные для контроля существующих активов в инфраструктуре и выявления новых;

– системы веб-фильтрации, предоставляющие данные о посещении сотрудниками подозрительных или запрещенных веб-сайтов [39, 177].

SIEM-анализ часто основан практически на «чистой» математике и статистике. Но отправной точкой служат задаваемые вручную правила. К примеру, однократное событие «*login failed*» не является существенным, в то время как пять и более таких событий для одной учетной записи уже могут

свидетельствовать о попытках подбора пароля. В простейшем случае, в SIEM-системах правила представлены в формате RBR (Rule-Based Reasoning) и содержат набор условий, триггеры, счетчики, сценарии действий. Корреляционный анализ позволяет выявлять неочевидные взаимосвязи между событиями ИБ.

SIEM-системы находят применение для решения задач мониторинга и обеспечения ИБ в контексте IoT. Так, в [104] предлагается система обнаружения распределенных DoS-атак (DDoS-атак) IoT-ботнетами на основе SIEM. Система обнаруживает и блокирует трафик DDoS-атаки от скомпрометированного IoT-устройства посредством мониторинга специфичных типов пакетов, включая пакеты TCP SYN, ICMP и DNS, исходящие от этих устройств. Показано, что подход, основанный на SIEM, может быть успешно сконфигурирован для точной идентификации и блокировки зловредного трафика скомпрометированных IoT-устройств. Технологии SIEM применяются для обнаружения инцидентов информационной безопасности IoT-систем также в [57]

В [159] предлагается интеграция SIEM-системы OSSIM с беспроводной системой обнаружения вторжений для повышения защищенности медицинских систем IoT. Система обнаружения вторжений разворачивается на Rasberry Pi, анализируется возможность использования такого устройства в качестве хоста для беспроводной IDS. Используются также возможности корреляции SIEM для сообщений об обнаруженных аномалиях и фильтрации ложных срабатываний. Результаты показали высокую эффективность подхода, даже в сильно загруженной среде предложенный подход позволил успешно анализировать траффик, загрузка процессора не превышала 5%, оперативной памяти – 400 МБ за 8 часов.

Одним из недостатков SIEM-системы является ее реактивная природа. По данным [22], проблема заключается в том, что SIEM-система начинает работать только тогда, когда злоумышленник уже проник в инфраструктуру. Поэтому для эффективной работы SOC-центра классические SIEM-системы необходимо

дополнять интеллектуальными системами, которые позволят обнаружить злоумышленника ещё на ранних стадиях атаки или на этапе подготовки к взлому.

Таким образом, подход на основе SIEM-систем, адаптированный для работы с большим количеством источников данных, хорошо подходит для интеграции в IIoT-системы с целью анализа и выявления инцидентов ИБ, но обязательно должен быть дополнен средствами, позволяющими обнаруживать злонамеренные действия на ранних этапах.

В [66] рассматривается архитектура безопасности IoT-систем, которая основывается на защите систем связи, устройств и взаимодействия в сети. Для защиты каналов применяются шифрование и проверка подлинности. Защита устройств рассматривается как обеспечение безопасности и целостности программного кода. Тема безопасности кода в [66] не обсуждается, целостность обеспечивается подписанием кода и проверкой подписи перед запуском. Есть вероятность изменения кода после запуска в момент загрузки, это компенсируется хостовыми средствами защиты, такими как харденинг (Hardening – усиление защищенности системы), разграничение доступа к системным файлам и ресурсам, контроль подключений и т.д. Контроль взаимодействия в сети основывается на аналитике, которая помогает выявить подозрительные и злонамеренные аномалии, угрозы, преодолевшие имеющиеся средства и системы защиты.

В работе [67], являющейся продолжением работы [66], подчеркивается, что средства мониторинга и аналитики могут быть единственным средством решения задач обеспечения безопасности в системах, где обновление приборов в промышленных системах управления невозможно без замены всей системы целиком (промышленное производство, нефтедобыча и пр.).

Подчеркивается, что в таких случаях системы обнаружения аномалий особенно полезны, что многие сети IoT характеризуются определенными шаблонами поведения, и отклонения легко идентифицируются. Дело усложняется широким набором протоколов, но на помощь приходит применение средств машинного обучения.

Остановимся подробнее на проблеме обнаружения атак и аномалий сетевого трафика промышленного Интернета вещей. В рассматриваемой трёхуровневой архитектуре IIoT (Рисунок 1.2) они должны выявляться на граничном уровне, включающем устройства IIoT, сети ближнего действия и пр.

Стоит отметить, что в работе не рассматриваются методы обеспечения безопасности облачных вычислений, широкой используемых в IIoT. Вопросы обеспечения их безопасности рассматриваются в [60].

Рассмотрим сетевые атаки подробнее. Согласно ГОСТ Р МЭК 56205-2014 [3], атака представляет собой умышленное посягательство на систему, продуманную попытку обойти сервисы безопасности и нарушить политику безопасности системы. Стандарт выделяет следующие типы атак:

– активная, предполагающая воздействие на ресурсы системы или ее работу;

– пассивная, нацеленная на получение информации без воздействия на систему и ресурсы;

– внутренняя, инициированная в пределах периметра безопасности;

– внешняя, инициированная за пределами периметра безопасности.

Численная оценка вероятности атаки и оперативное управление защитой информации на её основе подробно рассмотрены в [38]. На сегодняшний день можно выделить следующие виды сетевых атак на промышленные сети, представленные в Таблице 1.1 [24, 35, 49].

Таблица 1.1 – Виды сетевых атак и методы противодействия им

Вид атаки	Описание атаки	Реализация	Методы противодействия
IP-спуфинг	выдача злоумышленником себя в качестве легитимного пользователя посредством подмены IP адреса	внедрение ложной информации или вредоносных команд в стандартный поток данных	– использование криптографических средств аутентификации; – контроль доступа

Продолжение Таблицы 1.1

Вид атаки	Описание атаки	Реализация	Методы противодействия
Инъекции	межсайтовый скриптинг (XSS-атака), SQL-инъекция, XPath-инъекция.	модификация запроса к базе данных, внедрение в веб-страницу произвольного кода	– кодирование данных и управляющих символов; – правила построения SQL-запросов; – регулярное обновление.
Отказ в обслуживании (DoS)	нарушение доступности сервисов и систем посредством большого числа запросов	поддержание всех соединений в занятом состоянии	– функции анти-DoS; – функции анти-спуфинга; – применение систем обнаружения атак.
Фишинг-атаки	социальная разработка или обман сотрудников в целях хищения идентификацион-ных данных и дальнейшего несанкционирова нного использования	рассылка писем через электронную почту или сообщений в мессенджерах, как правило, содержащих ссылку на фишинговый ресурс, применение социальной инженерии	– использование проверенных ресурсов; – применение средств антивирусной защиты (САВЗ), в том числе почтовых, со своевременным обновлением базы сигнатур; – обучение и подготовка сотрудников.
Переполнение буфера (buffer overflow)	поиск уязвимостей, эксплуатация которых способна вызвать нарушение границ памяти, выполнить произвольный бинарный код от имени легитимного пользователя	1) подготовка кода для привилегирован-ного выполнения; 2) модификация последовательност и команд в программе для передачи управления подготовленному коду	– корректировки исходных кодов программы; – использование неисполнимых буферов; – применение проверок выхода за границы; – проведение проверок целостности.
Сетевая разведка	сбор информации о сети для планирования атаки	сетевая разведка посредством DNS- или ICMP-запросов (эхо), сканирования портов	– блокировка ICMP эхо ответов пограничных маршрутизаторов. – применение систем обнаружения атак.
Специализированные программы	вирусы, сетевые черви, троянский конь, сниффер, руткит	сбор данных скрытым нелегитимным агентом в системе, лавинообразное распространение	применение: – САВЗ с регулярно обновляемыми базами сигнатур; – шифрования; – антиснифферов; – межсетевых экранов; – антируткитов.

Актуальной является проблема выявления неизвестных сетевых атак (Zero-day), вероятность реализации которых предположительно можно снизить, используя предупреждающие меры. Инструментальные средства оценки уязвимостей рассматриваются в [23], методы оценки рисков реализации неизвестных сетевых атак – в [192], повышения эффективности их обнаружения в [18]. Также могут использоваться методы проактивного мониторинга, к примеру, на основе анализа временных рядов [40], а также системы обнаружения аномалий, в том числе на основе искусственных иммунных систем [148].

Под аномалией сетевого трафика понимается существенное отклонение трафика сетевого устройства от нормального профиля трафика для данного устройства или группы устройств [63].

Среди причин возникновения аномалий сетевого трафика можно выделить следующие [16]:

- неисправность сетевого оборудования;
- неверная работа приложений;
- случайные или преднамеренные действия со стороны легитимных пользователей;
- действия злоумышленников и т.д.

Среди методов обнаружения аномалий можно выделить статистические методы, которые строятся на собранной статистике параметров нормального сетевого трафика и сравнения с ней анализируемого трафика [71]. К примеру, может анализироваться объем передаваемых данных, имеющих общие характеристики, или число соединений за 5 минут. Если значения параметров резко отклоняются от ожидаемых, значит, возникла аномалия. Другим примером анализируемой характеристики может быть количество пакетов, приходящих на определенный порт [75]. Статистические методы выявления аномалий также рассматриваются в [65], подчеркивается, что чаще всего в коммерческих системах обнаружения аномалий сочетается использование статистики с методами машинного обучения.

На основании проведенного анализа можно сделать следующие выводы. Существующая нормативно-правовая база РФ затрагивает вопросы ИБ IIoT пока только косвенно. Вопросы регулирования обеспечения ИБ IIoT в определенной степени закрываются существующими нормативно-техническими документами, разработанными для информационных систем, АСУ ТП, объектов КИИ и др. Проблема обнаружения атак и аномалий сетевого трафика является одной из ключевых для IIoT-систем, ее решению должно уделяться первостепенное внимание. Подход к обнаружению инцидентов ИБ на основе SIEM-систем хорошо подходит для интеграции в IIoT-системы, подобное объединение должно повысить эффективность решения задач мониторинга ИБ промышленного Интернета вещей.

1.3 Методы мониторинга ИБ сетей промышленного Интернета вещей с применением технологий интеллектуального анализа данных

Мониторинг ИБ сетей IIoT осуществляется на основе данных о сетевом трафике. Задача сбора этих входных данных вещей усложняется использованием IIoT-устройствами различных протоколов и типов подключений. В системах Интернета вещей применяются:

– беспроводные локальные сети (Wireless Local Area Network, WLAN), беспроводные персональные сети (Wireless Personal Area Network, WPAN), включая сети ближнего (малого и среднего) радиуса действия, такие протоколы, как: Wi-Fi, 6LoWPAN, ZigBee IP, Thread, Z-Wave, ZigBee, WirelessHart, BLE 4.2 (Bluetooth Mesh), MiWi.

– энергоэффективные глобальные сети (Low-Power Wide Area Network, LPWAN), технологии для передачи небольших данных на дальние расстояния: LoRaWAN, SIGFOX, CIoT, 4G LTE, 5G, NB-IoT и другие [54].

При использовании мобильных устройств с прямым доступом в Интернет через SIM-карту, перехват трафика возможен с помощью способов, рассмотренных в [45], включая установку агента на SIM-карту или на само IoT-

устройство. В данной работе не рассматриваются вопросы сбора данных сетевого трафика IoT-устройств, использующих мобильный Интернет непосредственно. Сбор данных с относительно стационарных устройств, подключенных к сети Интернет посредством определенной внутренней сетевой инфраструктуры, осуществляется посредством сниффинга или зеркалирования трафика с сетевого оборудования.

Необходимо учитывать распределенный характер объекта мониторинга ИБ, при этом в качестве источников входных данных должны выступать не только устройства IIoT, но и сетевое оборудование: маршрутизаторы, межсетевые экраны (МЭ) и пр., имеется возможность взаимодействия с SIEM-системой. Также следует учитывать необходимость реализации пространственно-временной модели сбора входных данных для систем мониторинга. Входные данные должны иметь привязку к конкретным узлам сети IIoT и времени их регистрации, то есть являться темпоральными.

Согласно Рисунку 1.4, сетевое взаимодействие относится к граничному уровню архитектуры IIoT, а системы управления, диагностики и мониторинга – к уровню платформы. В целях мониторинга ИБ сети IIoT с граничного уровня могут быть собраны данные о сетевом взаимодействии, состоянии IIoT-устройств, с уровня платформы – данные о текущем состоянии сетей и конечных точек IIoT, общем количестве инцидентов, поступающие от внешних систем мониторинга. К таким системам могут относиться системы SIEM и SCADA, а также, к примеру, система обнаружения опасных состояний промышленных объектов, рассмотренная в [79], и пр. Предполагается сбор данных о сетевом взаимодействии с канального по транспортный уровней сетевой модели OSI.

Мониторинг ИБ сети IIoT, в первую очередь, предполагает анализ состояния сетевого трафика IIoT, информация о котором включает:

– временные ряды технологических параметров (ВРТП), то есть параметры (данные), обрабатываемые с помощью мультисенсорных сетей;

– внутренний сетевой трафик IIoT, то есть данные, передаваемые по каналам связи на каждом из уровней управления IIoT и между уровнями

управления, т.е., в терминологии серии стандартов ГОСТ Р 62443, трафик трактов;

– внешний сетевой трафик IIoT, т.е. данные, поступающие из внешней среды (Интернет, передатчики, провайдеры и т.д.) и передаваемые во внешнюю среду;

– данные, поступающие от взаимодействующей SIEM-системы, о событиях (инцидентах) ИБ.

Таким образом, система мониторинга ИБ сети IIoT должна быть распределенной, учитывать характер собираемых входных данных, гетерогенность соответствующих источников, что относится и к ИИС как нижнему уровню системы мониторинга ИБ сети. Необходимо учитывать разнородность входных данных также на этапе их нормализации (приведения к единому формату представления).

Отметим также, что определение конкретного состава собираемых и анализируемых данных зависит от используемых протоколов и технологий конкретного объекта. В данной работе не предлагается какой-либо определенный состав параметров, наиболее универсальный для всех сетей IIoT или, наоборот, наиболее подходящий для определенной узкой области. Для обучения и работы системы обнаружения атак может быть использован любой набор параметров, достаточный для определения на его основе безопасности того или иного сетевого взаимодействия, выбранный экспертами в соответствии с используемыми на конкретном объекте сетевыми технологиями, протоколами, обеспечивающий возможность эффективной классификации.

Вместе с тем, общий подход к нормализации данных должен включать в себя кодирование их качественных, текстовых или лингвистических значений числовыми параметрами, преобразование исходного диапазона количественных значений к используемому системой диапазону, подробнее вопросы нормализации рассматриваются в [76, 84].

На этапе обучения и тестирования СОА применительно к сети IIoT будем отталкиваться от параметров, используемых в различных, наиболее часто

используемых наборах данных о сетевых соединениях, содержащих параметры трафика как для нормальных сетевых соединений, так и для различного рода атак – датасетах (ДС). Данные по некоторым из них приведены в Таблице 1.2.

Таблица 1.2 – ДС, используемые для обучения СОА на IoT/IIoT

Наименование ДС	Кол-во параметров / атрибутов	Кол-во видов / классов атак	Специфика	Атаки
KDD-99 [142]	41	22	трафик информационно-телекоммуникационных сетей (ИТКС)	сетевые атаки на ИТКС
NSL-KDD [161]	41	22		
UNSW-NB15 [184]	48	9		
LITNET-2020 [123]	85	12		
TON_IoT [183]	44	9	Логи ОС Windows, Ubuntu, IoT/IIoT, сетевой трафик ИТКС	
AWID2 [115]	154	16	WiFi-трафик	сетевые атаки в WiFi-сетях
AWID3 [116]	253	13		
VARIoT [124]	83	0	трафик IoT-устройств умного дома	нет
IoTID20 [187]	83	4		атаки на IoT
IoT-23 [113,131]	21	8		ботнет-трафик
N-BaIoT [154]	115	2		
BoTNeTIoT-L01 [138]	23	2		
Bot-IoT [182]	46 или 10	4	трафик IoT-устройств	
NF-BoT-IoT [160]	43 или 12	4	NetFlow-версия Bot-IoT	
DS2OS [127]	12	7	данные прикладного уровня	
OTIDS [118]	4	3	данные CAN (Controller Area Network – сети контроллеров)	атаки на CAN
WSN-DS [106]	23	4	трафик WSN по протоколу LEACH	атаки на WSN
WUSTL-IIOT-2021 [194]	43	4	трафик IIoT по протоколу Modbus	атаки на IIoT

Примечание. В Таблице 2.2 для некоторых датасетов встречаются два значения количества параметров, к примеру, для Bot-IoT – «46 или 10». Имеется в виду, что существует две версии ДС с большим и меньшим количеством параметров. Кроме того, для датасета VARIoT указано количество содержащихся

атак, равное нулю; это связано с тем, что VARIoT содержит данные только о нормальном сетевом взаимодействии.

Также отметим, NF-BoT-IoT является NetFlow-версией датасета Bot-IoT, представленной Университетом Квинсленда (Австралия). На соответствующей странице [152] сайта данного университета представлены NetFlow-версии и других сетевых наборов данных, подробно описанные в [173], такие как NF-UNSW-NB15, NF-ToN-IoT, NF-BoT-IoT, NF-CSE-CIC-IDS2018, NF-UQ-NIDS.

BoTNeTIoT-L01 является доработанной версией N-BaIoT с уменьшенной избыточностью, с выбором только параметров 10-секундного временного окна. NSL-KDD – доработанная версия KDD-99, не содержащая избыточных и повторяющихся записей.

Большинство из этих ДС содержат или ботнет-трафик, т.е. трафик устройств, которые уже заражены, или трафик, характерный для классических информационно-телекоммуникационных систем (ИТКС), не содержащих данные, характерные для IIoT-устройств, или узкоспециализированные данные, такие как атрибуты CAN и данные прикладного уровня.

Как правило, классическими для построения, обучения и тестирования систем обнаружения вторжений считаются наборы обучающих данных (датасеты) KDD-99 и его усовершенствованная версия – NSL-KDD. Рассмотрим задачу нормализации параметров и уменьшения размерности пространства этих параметров на примере NSL-KDD.

NSL-KDD – содержит набор векторов-строк, каждая из которых состоит из 41 параметра соединения, представленных в Таблице 1.3. Каждая строка отмечена, соответствует ли она какому-либо виду атаки или нормальному состоянию системы. Всего представлено 22 класса атак, объединенных в 4 группы: User to Root (U2R), Remote to Local (R2L), Probe, Denial of Service (DoS). Все данные, представленные в наборе, могут быть сгруппированы в три категории: характеристики отдельных соединений, особенности соединения, параметры соединения за промежуток времени.

Таблица 1.3 – Параметры датасета NSL-KDD

Наименования параметров		
duration;	su_attempted;	same_srv_rate;
protocol_type;	num_root;	diff_srv_rate;
service;	num_file_creations;	srv_diff_host_rate;
flag;	num_shells;	dst_host_count;
src_bytes;	num_access_files;	dst_host_srv_count;
dst_bytes;	num_outbound_cmds;	dst_host_same_srv_rate;
land;	is_host_login;	dst_host_diff_srv_rate;
wrong_fragment;	is_guest_login;	dst_host_same_src_port_rate;
urgent;	count;	dst_host_srv_diff_host_rate;
hot;	srv_count;	dst_host_serror_rate;
num_failed_logins;	serror_rate;	dst_host_srv_serror_rate;
logged_in;	srv_serror_rate;	dst_host_rerror_rate;
num_compromised;	rerror_rate;	dst_host_srv_rerror_rate.
root_shell;	srv_rerror_rate;	

Классы сетевых атак, представленные в NSL-KDD, приведены в Таблице 1.4.

Таблица 1.4 – Сетевые атаки, содержащиеся в NSL-KDD

Вид атаки	Класс атаки
R2L	ftp_write
	guess_passwd
	imap
	multihop
	phf
	spy
	warezclient
	warezmaster
probe	ipsweep
	nmap
	portsweep
	satan
DoS	Back
	land
	neptune
	pod
	smurf
	teardrop

Продолжение таблицы 1.4

Вид атаки	Класс атаки
U2R	buffer_overflow
	loadmodule
	perl
	rootkit

В [26, 53, 62, 83, 121, 157, 158, 198] на примере датасета KDD-99, который по составу используемых параметров (признаков) сетевого трафика аналогичен рассматриваемому NSL-KDD, предлагаются различные подходы к уменьшению размерности (сжатию) пространства признаков, т.е. к определению наиболее информативных из них. В [53] для сжатия пространства параметров применяется сингулярное разложение матриц. В [26] решается задача сжатия пространства параметров NSL-KDD с использованием методики расчета влияния атрибутов через механизм анализа соответствий.

Однако в [83, 121, 157, 158, 198] предлагается для каждой атаки использовать свой отдельный список параметров. То есть для атаки DoS предлагается один набор параметров, для U2R – другой и т.д. При использовании параметров, предлагаемых в [62], в вычислительных экспериментах обнаружить угрозы вообще не удалось. Сжатие с использованием сингулярного разложения матриц позволяет проводить классификацию единовременно сжатых данных, однако не позволяет анализировать все новые и новые строки сетевой активности. Поэтому в данной работе решалась также задача определения информативных (значимых) параметров используемых датасетов, подробнее рассматриваемая в разделе 2.2.

После сбора и нормализации данные о сетевом трафике подвергаются анализу. Для этого используются различные методы и технологии интеллектуального анализа данных, подробно рассмотренные в [125, 175]. Интеллектуальные системы сегодня в первую очередь ассоциируются с искусственными нейронными сетями (ИНС), зарекомендовавшими себя в качестве эффективного метода классификации, применяющегося для решения огромного количества различных задач, включая обнаружение атак и сетевых

аномалий [43, 48, 83], в том числе с использованием технологий глубокого обучения [44, 82]. Методы машинного обучения также включают в себя алгоритмы деревьев решений (ДР), случайного леса (СЛ), муравьиной колонии, нечеткой логики, k-ближайших соседей (k-nearest neighbors, KNN), машины опорных векторов (SVM), наивный байесовский классификатор (НБК), генетические алгоритмы (ГА) и др. Эти методы широко используются для решения задач интеллектуального анализа данных, в том числе для выявления и классификации атак [21, 37, 52], на основе анализа данных о сетевом трафике.

Для решения задач обнаружения атак также нашли достаточно широкое применение искусственные иммунные системы (ИИС) [59]. Они характеризуются способностью обнаруживать неизвестные атаки, возможностью постоянного фонового самообучения. По данным [179-180], ИИС могут в разы превосходить своих главных конкурентов (ИНС и ГА) по быстродействию, а также характеризуются вдвое меньшим количеством ошибок. Другим преимуществом ИИС является их применимость в распределенных системах, в том числе в рамках реализации многоагентного подхода [151].

Повышение эффективности обнаружения сетевых атак и аномалий возможно также за счет совместного использования нескольких методов искусственного интеллекта (ИИ) в рамках гибридной интеллектуальной системы (ГИС), объединяющей в своем составе две или более различные технологии ИИ с целью получения синергетического эффекта, нивелирования недостатков одной технологии преимуществами другой. Так, системы нечеткой логики понятны и прозрачны для пользователя, но у них отсутствует способность к обучению. ИНС, наоборот, способны к обучению, но непрозрачны для пользователя. Их совместное использование в составе нечеткой нейронной сети (ННС) позволяет получить адаптивную систему, способную к обучению и одновременно в значительной степени прозрачную для пользователя [136].

Общая идея построения гибридных интеллектуальных систем обнаружения атак (ГИСОА) обсуждается в ряде работ [105, 119, 143, 167, 168, 172]. Как правило, в основе построения ГИСОА используется объединение ИНС,

алгоритмов кластерного анализа, ДР, SVM с другими различными по своей идеологии методами ИИ. Отдельную перспективную группу СОА занимают ГИСОА на базе ИИС в дополнении с другими технологиями ИИ [77, 133, 137, 144-145, 153, 169, 188].

Применительно к промышленным сетям и сетям Интернета вещей используются подробно рассмотренные в [162] подходы, включающие в себя:

– нейронные сети глубокого обучения – для обнаружения сетевых атак и аномалий в IoT/IIoT [36, 82, 149, 156, 197], сетевых атак на киберфизические системы [78];

– методы машинного обучения – для обнаружения атак на IoT [81, 156], IIoT [31, 109, 110, 126, 150], в том числе в сравнении различных алгоритмов с ИНС [110, 150], где лучшую точность демонстрируют ИНС и СЛ;

– искусственные иммунные системы – для идентификации вторжений в сети IoT [103], IIoT [102, 195];

– гибридные интеллектуальные системы обнаружения атак на IoT [117, 128, 134, 171], IoMT [147], АСУ ТП [17, 20] и IIoT [101, 112, 132, 170], в том числе использующие алгоритмы ИИС [170].

Таким образом, для решения задачи обеспечения сетевой безопасности промышленного Интернета вещей сегодня предложены различные методы ИИ, включая методы машинного обучения, ИНС, ИИС, а также гибридные интеллектуальные системы, позволяющие использовать преимущества различных подходов и нивелировать их недостатки. Особый интерес в данном случае представляет применение ИИС, в том числе в составе гибридных интеллектуальных систем, благодаря их способности обнаруживать неизвестные атаки, постоянно самообучаться, высокому быстродействию, низкому уровню ошибок, применимости для реализации в классе распределённых многоагентных систем.

В данной работе предлагается соответствующая концепция построения многоуровневой гибридной распределенной интеллектуальной системы обнаружения атак и аномалий сетевого трафика IIoT на основе многоагентной

платформы с использованием механизмов искусственных иммунных систем, методов машинного обучения, взаимодействия с подсистемой корреляционного анализа событий ИБ или SIEM-системой, подробно изложенной в разделе 3.1.

1.4 Применение искусственных иммунных систем для решения задачи обнаружения атак и аномалий сетевого трафика

1.4.1 Основные алгоритмы теории искусственных иммунных систем

Для лучшего понимания алгоритмов ИИС, кратко рассмотрим работу естественной иммунной системы (ЕИС), служащей для защиты организма человека от чужеродных зловредных организмов – патогенов. В ЕИС выделяют две значимые подсистемы: врожденный иммунитет, приобретенный иммунитет. Первый передаётся по наследству, не меняется в течение жизни. Его механизмы обеспечивают первую линию защиты, включают механические барьеры (кожа, слизистые оболочки), гуморальные факторы (цитокины, система комплемента и пр.), клеточные механизмы. Второй осуществляет специфичный иммунный ответ на конкретный вид распознанного патогена с помощью специальных клеток – лимфоцитов. В ИИС чаще всего моделируются механизмы функционирования лимфоцитов, распознающих и атакующих «чужого», и дендритных клеток (ДК), осуществляющих анализ уровня опасности по объему повреждений тканей, запускающих или угнетающих деятельность лимфоцитов.

Построение искусственной системы, обладающей всеми полезными свойствами и функционалом ЕИС, является пока нереализуемой задачей, но существуют алгоритмы, успешно имитирующие некоторые функции ЕИС, позволяющие решать задачи в том числе обнаружения сетевых атак и аномалий.

Алгоритм негативной селекции (Negative Selection Algorithm, NSA) нацелен на выполнение классификации «своего» и «чужого». Предполагает в первую очередь определение обучающих примеров «своего» (S) как совокупности строк длиной l, затем формирование случайным образом детекторов (D) таких, что ни

один детектор не соответствует ни одной строке *S*. После чего анализируемые данные (*A)* приводятся к соответствующему строковому виду, и каждая строка *A* сравнивается с каждым детектором *D*. Если находится соответствие, то строка A_i считается «чужой».

Для определения факта соответствия используются различные варианты оценки мер близости между точками или векторами в пространстве анализируемых параметров:

– Евклидово расстояние:

$$d_E(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}; \quad (1.1)$$

– квадрат Евклидова расстояния:

$$d_{E^2}(x,y) = \sum_{i=1}^{n}(x_i - y_i)^2; \quad (1.2)$$

– Манхеттенское расстояние:

$$d_M(x,y) = \sum_{i=1}^{n}|x_i - y_i|; \quad (1.3)$$

– расстояние Хэмминга – предполагает подсчет количества координат, по которым векторы X и Y отличаются;

– косинусное расстояние:

$$\cos\varphi = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} {x_i}^2}\sqrt{\sum_{i=1}^{n} {y_i}^2}}; \quad (1.4)$$

– степенное расстояние:

$$d_C(x, y) = \sqrt[r]{\sum_{i=1}^{n}(x_i - y_i)^p}; \quad (1.5)$$

где r и p – параметры, значения которых выбираются пользователем; x_i и y_i – компоненты векторов параметров X и Y сравниваемых шаблонов детекторов; n – размерность этих векторов;

Недостатком применения алгоритма негативной селекции (NSA) в классическом виде, согласно [25], является тот факт, что при линейном увеличении количества обучающих данных о нормальном состоянии системы, необходимое количество детекторов для той же точности обнаружения атак увеличивается экспоненциально.

Модель гиперклетки представляет собой модификацию алгоритма NSA, при котором создается большая гиперклетка, покрывающая области как «своего», так и «чужого», затем от нее отсекаются части, покрывающие нецелевые области. Существуют различные вариации данного алгоритма: может изменяться форма гиперклетки, вид, критерий останова и пр. По данным [25], модель гиперклетки обладает более высокой скоростью обучения.

Алгоритм NSA обеспечивает высокую эффективность СОА, однако может быть заменен также набором экспертных правил, создаваемых с помощью дерева решений, как это предложено в [27].

Алгоритм клональной селекции (Clonal Selection Algorithm, CSA) имитирует процесс пролиферации активированного В-лимфоцита. Существуют различные подходы к его реализации. К примеру, в [25] описывается создание детекторов, обнаруживающих нормальное состояние системы. Для этого детекторы генерируются случайно, определяется аффинность a_i как мера близости каждого детектора $\vec{d}_i$ и вектора данных нормального состояния $\vec{s}_j$. Детекторы с наибольшим значением a клонируются в количестве, прямо пропорциональном a. Каждый клон подвергается мутации, степень мутации обратно пропорциональна a. Вероятно, на этапе анализа предполагается, что если ни один детектор не соответствует анализируемой строке, то она отмечается аномальной.

В [176], наоборот, детекторам представляют обучающие данные, соответствующие аномалиям A. В результате обучения получаются детекторы, соответствие с которыми свидетельствует о нештатном состоянии контролируемой системы или процесса. Здесь предполагается генерация детекторов каждого обнаруживаемого класса пропорционально количеству обучающих примеров соответствующего класса. Количество необходимых детекторов находится как:

$$|D_k| = \frac{|A_k|}{|A|} \times |D| \tag{1.6}$$

где $|D_k|$ – количество детекторов во множестве детекторов D, соответствующих k-му классу данных; $|A_k|$ – количество строк обучающих данных об аномалиях класса k; $|A|$ – общее количество обучающих строк данных об аномалиях; $|D|$ – количество детекторов во множестве детекторов D.

Разрастание количества детекторов в результате клонирования может негативно сказываться на производительности системы, поэтому часто ограничивают количество детекторов и срок их существования. Если срок существования детектора превышен, он уничтожается, вместо него генерируется новый. Если детектор обнаруживает аномалию, срок его существования значительно увеличивается.

В целом, алгоритм клональной селекции является адаптивным алгоритмом, позволяющим дообучать систему в процессе эксплуатации, но отдельно его применение не обеспечивает толерантности системы к «своим» данным.

В соответствии с теорией опасности, иммунная система при определении необходимости своего реагирования использует не только обнаружение чужеродных патогенов, но и в большей степени учитывает опасность той или иной ситуации. То есть иммунитет должен более агрессивно реагировать не на «чужое, но безопасное», а скорее на «свое, но опасное» [30].

С теорией опасности тесно связан алгоритм дендритных клеток (ДК). ДК аккумулируют сигналы опасности, имеют три состояния: незрелое, полузрелое, зрелое. По данным [41], незрелое состояние – начальное состояние ДК, находящихся в поиске патогенов, полузрелое состояние устанавливается в случае обнаружения безопасного патогена, зрелое – опасного. В [41] предлагается предусмотреть для ДК три входа: сигнал опасности, сигнал безопасности, сигнал о наличии патогена (Pathogen-Associated Molecular Patterns, PAMP) и три выхода: сигнал о костимуляции, сигнал о полузрелом состоянии, сигнал о зрелом состоянии. Для определения выхода предлагается общая формула:

$$Output = (P_\omega \sum_i P_i + D_\omega \sum_i D_i + S_\omega \sum_i S_i)(1 + I), \quad (1.7)$$

где $P_\omega, D_\omega, S_\omega$ – веса сигналов PAMP, опасности, безопасности соответственно, $\mathrm{P_i}, \mathrm{D_i}, \mathrm{S_i}$ – входные сигналы PAMP, опасности, безопасности соответственно, I – сигнал воспаления.

Хотя здесь не совсем понятно, для чего нужна операция суммирования, которая нейтрализует разделение видов сигналов. Авторы приводят цитату формулировки алгоритма ДК из [163], который при его упрощении заключается в следующем порядке действий:

1) создать множество ДК, выбрать из него случайным образом подмножество ДК;

2) представить каждую единицу анализируемых данных дендритным клеткам из выбранного подмножества, определить уровни входных сигналов, на основе (1.7) вычислить выходные сигналы;

3) если сигнал о зрелом состоянии выше, чем о полузрелом, отметить ДК как зрелую, иначе как полузрелую;

4) перевести ДК в новое множество дифференцированных ДК, если сигнал костимуляции превышает некоторый пороговый уровень, добавить новую ДК в анализирующее эту единицу данных подмножество.

После анализа данного алгоритма возникло предположение, что его авторы скорее имели в виду, что выходной сигнал является не суммой трех выходов, а их совокупностью – множеством:

$$Output = \{P_\omega \sum_i P_i\,(1+I), D_\omega \sum_i D_i\,(1+I), S_\omega \sum_i S_i\,(1+I)\}. \quad (1.8)$$

Заметим, что алгоритм ДК и теория опасности позволяют сконцентрировать реагирование больше на действительные (реальные) угрозы, чем на случайные аномалии.

Теория идиотипической иммунной сети основана на предположении о функционировании иммунной системы как регулирующей сети антител, взаимно стимулирующих друг друга даже в отсутствие патогенов. Согласно данной теории, иммунная система активирует сама себя, обеспечивая поддержание детекторов в активном состоянии даже в отсутствие антигена, имитируя его присутствие.

Предполагается, что лимфоциты, способные распознавать любые чужеродные антигены, должны распознавать и соответствующие им антитела, вырабатываемые другими лимфоцитами. В результате антитела одних лимфоцитов распознаются другими лимфоцитами, что поддерживает активированное состояние последних. В [86] подробно описано формирование идиотипической сети. Таким образом, механизмы идиотипической иммунной сети более актуальны для ЕИС, чем для ИИС, так как в ИИС для поддержания активности детектора достаточно программно задать соответствующие параметры, не нужно для этого имитировать вторжение.

В целом, сегодня существуют различные подходы к построению ИИС, каждый из них характеризуется определенными преимуществами и недостатками. Так, алгоритм негативной селекции характеризуется низким числом ошибок первого рода, способностью обнаруживать неизвестные аномалии, но низкой адаптивностью. Механизмы клональной селекции обеспечивают адаптивность системы, но не гарантируют достаточно низкого уровня ошибок первого рода.

Часто эти два селективных алгоритма используются в комплексе, такие системы демонстрируют высокую эффективность.

Алгоритмы дендритных клеток, сконцентрированные на сигналах опасности, требуют разработки корректных экспертных правил в определении того, что является опасностью. Более того, они не содержат какого-либо общего механизма выявления аномалии или нормы, а только определяют уровень опасности выявленных другими механизмами аномалий и нормы.

Теория идиотипических иммунных сетей, где детекторы имитируют аномалии для продления срока существования друг друга, легко заменяется простым программным увеличением срока их существования. Однако формирование иммунной сети, основанной на другом взаимодействии, когда один узел распределенной в пространстве системы выявил некоторую угрозу и передал информацию о ней всем другим узлам, является перспективным.

Объединение рассмотренных подходов к построению ИИС в единую распределенную систему обнаружения атак и аномалий позволяет получить синергетический эффект, при котором на основе негативной селекции возможно обнаружение в том числе неизвестных угроз при низком уровне ошибок первого рода, а наличие подсистемы клональной селекции обеспечивает адаптивность системы, что позволяет ей на основе выявленной неизвестной угрозы самообучиться лучшему выявлению подобных угроз. Механизмы обновления и замены детекторов обеспечивают при этом стабильный размер популяции, не допуская перегрузку ими системы.

Гибридизация алгоритмов негативной и клональной селекций сама по себе не нова, однако их дополнение модифицированной подсистемой дендритных клеток (ДК) позволяет обеспечить реагирование системы прямо пропорционально опасности. Предлагается модификация ДК-алгоритма, при которой ДК не накапливает сигналы безопасности, что в противном случае позволило бы злоумышленнику имитировать сигналы безопасности, выполняя активность, характерную для атаки, а ДК ассоциировала бы данную активность с сигналами безопасности и обеспечила толерантность системы к атаке, что в дальнейшем

позволило бы злоумышленнику проводить атаку, на которую система бы уже не реагировала. Напротив, ДК должна только оценивать уровень опасности: если он нулевой или близкий к нему, то контрмеры должны быть минимальными, не «параноидальными». Если подобные или связанные аномалии встречаются чаще, то уровень ассоциированной с ними опасности накапливается, и, соответственно, инициируются более весомые контрмеры. Результат применения такого объединения алгоритмов CSA и NSA с алгоритмом ДК для анализа датасета Bot-IoT опубликован в [30].

Построение подобной системы в архитектуре распределенных взаимосвязанных компонентов позволяет, в случае обнаружения неизвестной угрозы в одном сегменте сети, мгновенно обучить лучшему обнаружению подобных угроз все другие сегменты сети. Более того, единая подсистема ДК получает возможность оценивать сигналы опасности с разных сегментов, выстраивая общую картину безопасности.

Таким образом, особый интерес вызывает построение гибридной ИИС, объединяющей алгоритмы лимфоцитов и дендритных клеток в классе распределенной двухуровневой системы взаимодействующих агентов.

1.4.2 Применение искусственных иммунных систем для обеспечения безопасности сетей промышленного Интернета вещей

Известны примеры применения ИИС для решения задач обеспечения безопасности IIoT. Так, в [102] предлагается многоуровневая система обнаружения вторжений для беспроводных сетей (WSN) на основе иммунной теории. Система включает блоки *B*-клеток, *T*-клеток, дентритных клеток и базофилов. Здесь *B*-клетки проводят первичный анализ данных, они формируются только на этапе обучения системы. Для измерения расстояния между векторами используется алгоритм побитового сопоставления. Дальнейший анализ данных производится дентритными клетками, в случае выявления аномалии передается сигнал блоку Т-клеток, который формирует реакцию, изолирует аномальный узел

и не участвует в анализе. Блок базофилов в работе не реализован. Но описанная авторами система не способна постоянно обучаться, так как отсутствует реализация алгоритма клональной селекции, сравнения с использованием других мер расстояния между векторами не проведено.

В [103] предложен алгоритм глубокого обучения и дентритных клеток (Deep Learning and Dendritic Cell Algorithm, DeepDCA). Реализовано сжатие пространства параметров, применяется самоорганизующаяся ИНС, осуществляющая первичную обработку данных и категорирование входного сигнала на сигналы об опасности и о безопасном состоянии. Анализ опасности осуществляется дендритными клетками. Представлены результаты сравнения с такими классификаторами как k-ближайших соседей, машина опорных векторов, многослойный персептрон, НБК. DeepDCA продемонстрировал наилучшую точность обнаружения. Однако не раскрывается весь потенциал ИИС: в данной работе речь идет об определении опасности и обнаружении только известных атак на основе данных от ИНС, реализация алгоритма негативной селекции позволило бы обнаруживать неизвестные атаки.

В [195], как и в предлагаемом подходе, используются алгоритмы негативной селекции для обеспечения толерантности системы к нормальному состоянию, клональной селекции, обеспечивающей адаптивность системы, возможность ее постоянного самообучения. В моделировании использован протокол LEACH, проанализированы следующие виды атак: Resource depletion, Sinkhole, Wormhole, Sybil, Selective forwarding attack. Обнаружение строится с использованием теории опасности. В первую очередь, экземпляры и центры кластера в WSN обнаруживают изменения своих собственных свойств, извлекают ключевые данные и получают информацию о сигналах среды, оценивают риск.

В случае опасности, экземпляр кластера передает соответствующий сигнал опасности центру кластера, объединяющего несколько сигналов опасности, переводящих их узлу-приемнику. Узел-приемник вычисляет степень риска, область риска, запрашивает представления антигенов. Только потом узлы датчиков опасной зоны собирают информацию о сетевом трафике для

формирования антигенов, и после этого узел-приемник проводит анализ на предмет вторжения. Но в таком случае атака будет обнаружена только после того как какое-либо оборудование будет уже неправомерно выведено из штатного режима функционирования, и только если такое выведение из штатного режима функционирования система посчитает опасной. То есть система будет игнорировать в том числе любое сканирование сети, что недопустимо.

Как показывает анализ, сегодня остается открытым вопрос о построении двухуровневой ИИС для защиты сетей промышленного Интернета вещей, реализующей комплекс иммунных механизмов в классе распределенных многоагентных систем, одновременно являющейся:

– адаптивной: способной обнаруживать неизвестные атаки и аномалии и самообучаться на их основе по алгоритму клональной селекции;

– распределенной по подсетям, независимо от их удаленности, состоящей из множества взаимодействующих и взаимообучающих агентов;

– реализующей обучение посредством механизма генерации случайных детекторов и их фильтрации по алгоритму негативной селекции, обеспечивающему толерантность детекторов к нормальному состоянию контролируемой системы;

– включающей реализацию блока дендритных клеток, в соответствии с теорией опасности анализирующего критичность тех или иных событий, что позволяет избежать параноидального реагирования на каждую сетевую аномалию;

– использующей принципы децентрализованного взаимодействия агентов нижнего уровня и централизованного взаимодействия с агентом(ми) верхнего уровня;

– осуществляющей мониторинг сетевого трафика постоянно, а не только после того, как был осуществлено воздействие на какой-либо контролируемый узел.

Построение такой ИИС подробно рассматривается в разделе 2. Кроме того, особый интерес вызывает реализация предложенного варианта построения ИИС в

составе гибридной интеллектуальной системы, что рассматривается в разделе 3. Построение гибридных многоуровневых интеллектуальных систем обнаружения атак позволяет получить лучшие значения показателей эффективности обнаружения атак по сравнению с отдельными интеллектуальными системами [42]. Алгоритмы ИИС широко применяются для защиты сетей IIoT в таких гибридных система.

Так, в [170] предложена двухуровневая модель обнаружения атак на IIoT, объединяющая глубокое обучение и алгоритм негативной селекции. Для анализа использовались наборы данных CICIDS 2017, CICIDS 2018 и ToN-IoT. Авторами строится двухуровневая система обнаружения сетевых вторжений на основе глубокого обучения DL-TL-NIDS (Deep Learning-based Two Level Network Intrusion Detection System). Для первого уровня используемый датасет балансируется, затем на нём обучается и тестируется глубокая нейронная сеть. Второй уровень использует два классификатора: алгоритм негативной селекции и глубокую нейронную сеть, обученную с помощью алгоритма стрекозы (Dragonfly Algorithm). Для комбинирования выходов каждой модели используется правило Демпстера-Шафера. Результаты показывают высокую эффективность модели.

Другим примером реализации гибридного подхода с использованием алгоритмов ИИС для защиты от сетевых атак в гетерогенных сетях является рассмотренный ранее алгоритм DeepDCA, представленный в [103].

Совместное использование ИИС и самоорганизующейся карты Кохонена в [169] позволило повысить эффективность обнаружения атак DoS и U2R при низком уровне ошибок первого рода. В данном случае работа СОА происходит в 2 этапа:

1) фильтрация признаков сетевых соединений с помощью иммунных детекторов, обученных по методу отрицательного отбора; тем самым отсеиваются те образцы, которые соответствуют нормальным сетевым соединениям;

2) аномальные экземпляры обрабатываются самоорганизующимися картами Кохонена и группируются в отдельные кластеры со схожими признаками.

В [133, 145, 188] в качестве иммунных детекторов выбраны многослойные ИНС, которые генерируются при помощи метода клональной селекции. В [77] в роли детекторов используются ИНС Кохонена, реагирующие на изменение статистики сетевого трафика. Блок формирования иммунной памяти реализует операции клонирования и мутации детекторов; мутация заключается в случайном изменении весов ИНС-детектора на малую величину, механизм клонирования детекторов заключается в создании 5 копий детектора, обнаружившего аномалию.

В [137] предложена модель распределенной сетевой СОА, состоящей из автономных взаимодействующих друг с другом агентов, реализующей алгоритм негативной селекции ИИС в комбинации с генетическим алгоритмом (ГА) для генерации детекторов. Эффективность предложенной модели сравнивалась экспериментально с машиной опорных векторов (SVM), наивным байесовским классификатором (НБК) и классификатором дерева решений J48. Предложенная модель продемонстрировала наилучшую эффективность.

Таким образом, анализ литературы показывает, что применение гибридных интеллектуальных систем на основе объединения ИИС с другими методами машинного обучения позволяет существенно повысить значения показателей эффективности обнаружения сетевых атак и аномалий промышленного Интернета вещей, однако в существующих работах также используются лишь отдельные теории и алгоритмы искусственных иммунных систем, упущен синергетический эффект применения совокупности алгоритмов и теорий ИИС.

В целом, ИИС является распределенной адаптивной системой обнаружения событий ИБ, нехарактерных для нормального состояния системы, или, в зависимости от реализации, опасных событий, в том числе неизвестных. Существуют примеры применения ИИС для защиты сетей промышленного Интернета вещей, однако открытыми остаются вопросы построения перспективной двухуровневой ИИС, основанной на комплексировании алгоритмов негативной селекции, клонального отбора, дендритных клеток, обновления и памяти детекторов, реализующей взаимное обучение и самообучение агентов нижнего уровня, анализ опасности на верхнем уровне в

составе гибридной распределенной системы мониторинга ИБ сетей промышленного Интернета вещей.

Выводы по первой главе

1. Системы промышленного Интернета вещей в настоящее время являются в достаточной мере уязвимыми. Система стандартизации в данной новой области пока только разрабатывается, производители все еще ставят в приоритет выгоду производства и функциональность устройств в ущерб их безопасности. Вместе с тем, уже сейчас начинают активно разрабатываться методы и средства защиты IIoT.

2. Существующая нормативно-правовая база РФ в области ИБ затрагивает вопросы обеспечения безопасности IIoT преимущественно только косвенно. Нормативно-техническая база обеспечения ИБ IIoT в определенной степени закрывается существующими нормативно-техническими документами, разработанными для информационных систем, АСУ ТП, значимых объектов КИИ и др.

3. В целях обеспечения сетевой безопасности промышленного Интернета вещей находят применение различные методы искусственного интеллекта, включая методы машинного обучения, ИНС, а также гибридные интеллектуальные системы, позволяющие использовать преимущества различных подходов и нивелировать их недостатки. Особый интерес представляет использование ИИС, в том числе в составе гибридных интеллектуальных систем, благодаря их способности обнаруживать неизвестные атаки, самообучаться, высокому быстродействию, низкому числу ошибок, применимости для реализации в классе распределённых многоагентных систем.

4. ИИС является распределенной адаптивной системой обнаружения событий ИБ, нехарактерных для нормального состояния системы, или, в зависимости от реализации, опасных событий, в том числе неизвестных. Существуют примеры применения ИИС для защиты сетей промышленного

Интернета вещей, однако открытыми остаются вопросы построения перспективной двухуровневой ИИС, основанной на комплексировании алгоритмов негативной селекции, клонального отбора, дендритных клеток, обновления и памяти детекторов, реализующей взаимное обучение и самообучение агентов нижнего уровня, анализ опасности на верхнем уровне в составе гибридной распределенной системы мониторинга ИБ сетей промышленного Интернета вещей.

2 Разработка и исследование алгоритмов обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием адаптивных механизмов искусственных иммунных систем

2.1 Функциональная модель процесса мониторинга информационной безопасности сети промышленного Интернета вещей

Решение задачи мониторинга ИБ систем IIoT усложняется наличием и использованием различных сетевых протоколов и технологий, однако в целом в процессе мониторинга ИБ решаются такие общие базовые задачи, как: перехват сетевого трафика, его анализ, принятие решения о наличии и классе атаки (или ее отсутствии), протоколирование, оповещение.

Для решения рассматриваемого круга задач в работе предлагается многоуровневая схема интеллектуального анализа данных трафика, где на нижних двух уровнях используется распределенная двухуровневая искусственная иммунная система (ИИС), на верхнем – система классификации состояния сетевого трафика IIoT. Функциональная модель процесса мониторинга ИБ сети IIoT представлена на Рисунке 2.1.

Рассмотрим подробнее функционирование предлагаемой ИИС. В первую очередь, эта система должна получать входные данные о сетевом трафике. Для этого они должны быть перехвачены или получены от сетевого оборудования, затем необходимо выделить анализируемые параметры (признаки) и привести их к определенному виду (нормализация).

ИИС располагает множеством агентов нижнего уровня, распределённых по подсетям, содержащих детекторы (искусственные лимфоциты), функционирующие по принципу «свой/чужой», выявляющие атаки, в том числе неизвестные и классифицирующие известные.

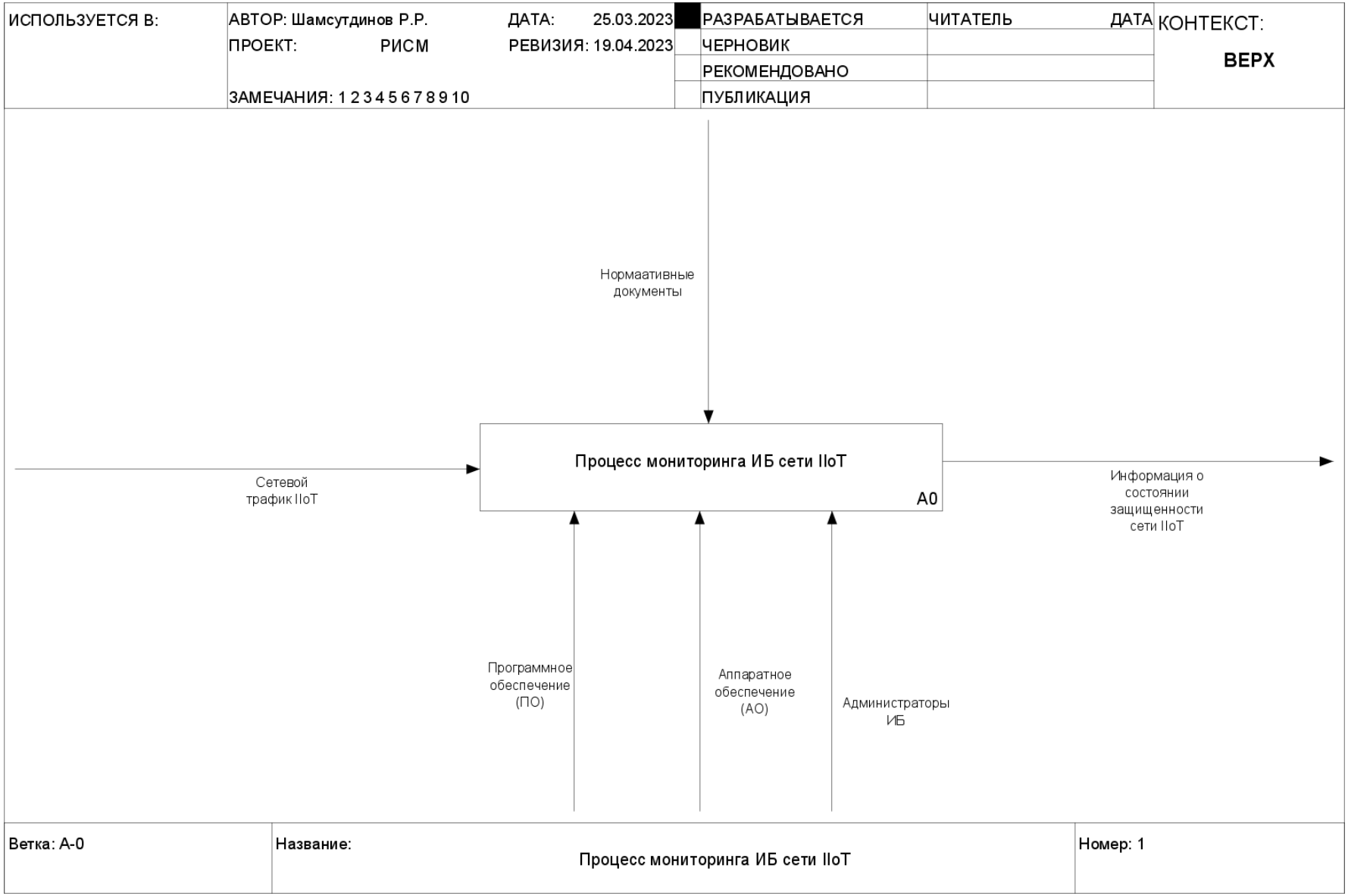


Рисунок 2.1 – Функциональная модель процесса мониторинга ИБ сети IIoT (лист 1 из 6)

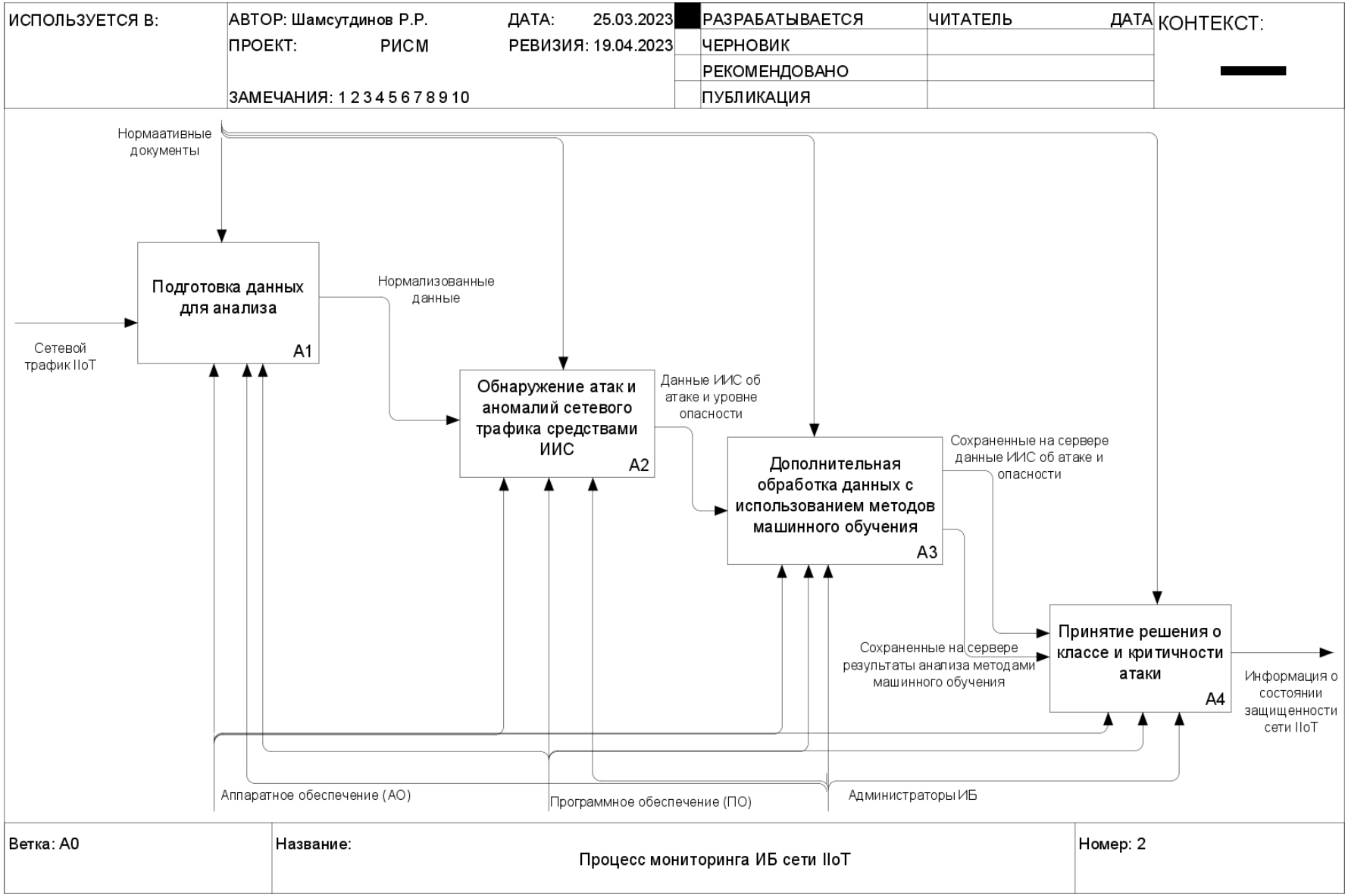


Рисунок 2.1 (лист 2 из 6)

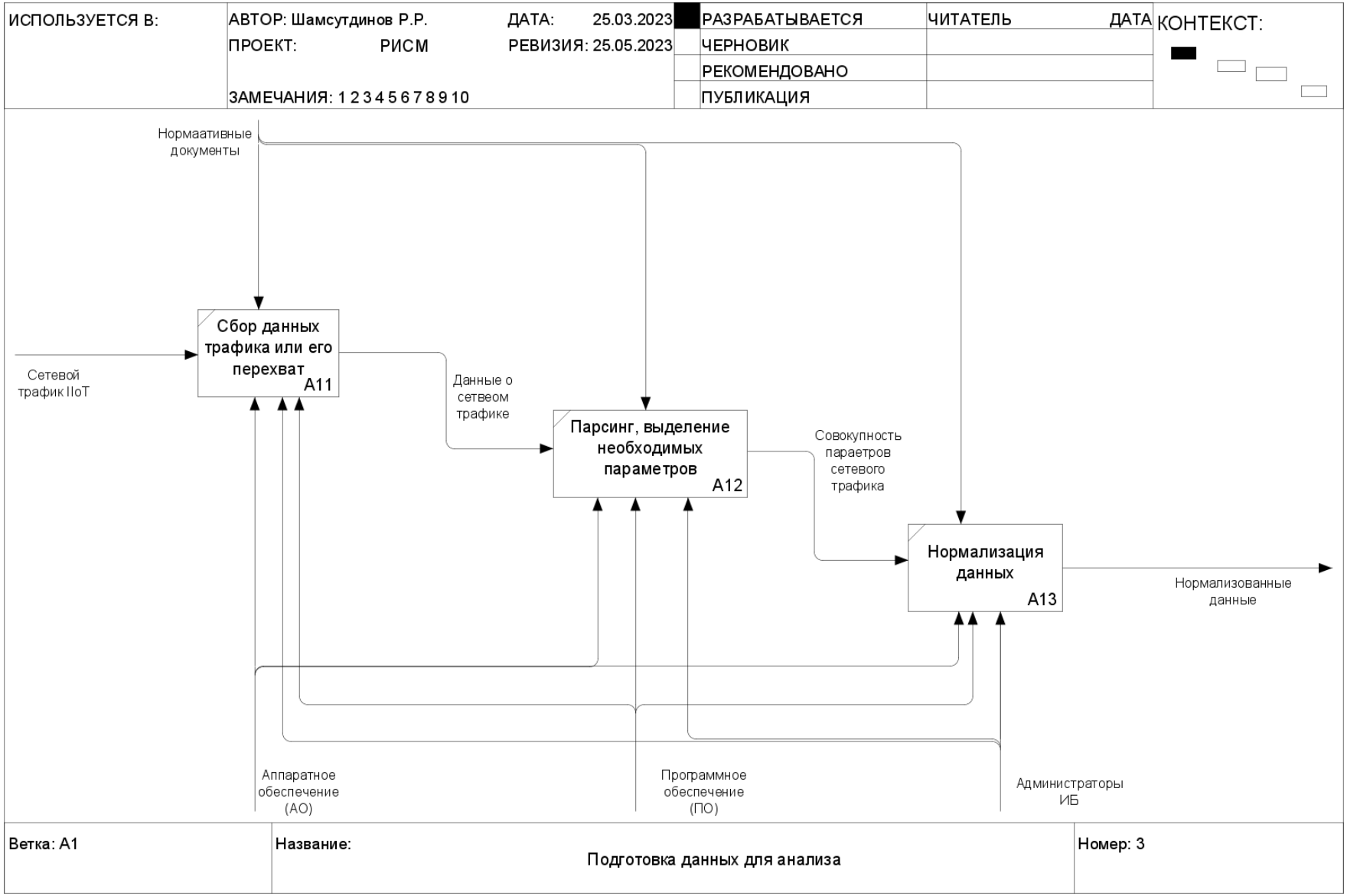


Рисунок 2.1 (лист 3 из 6)

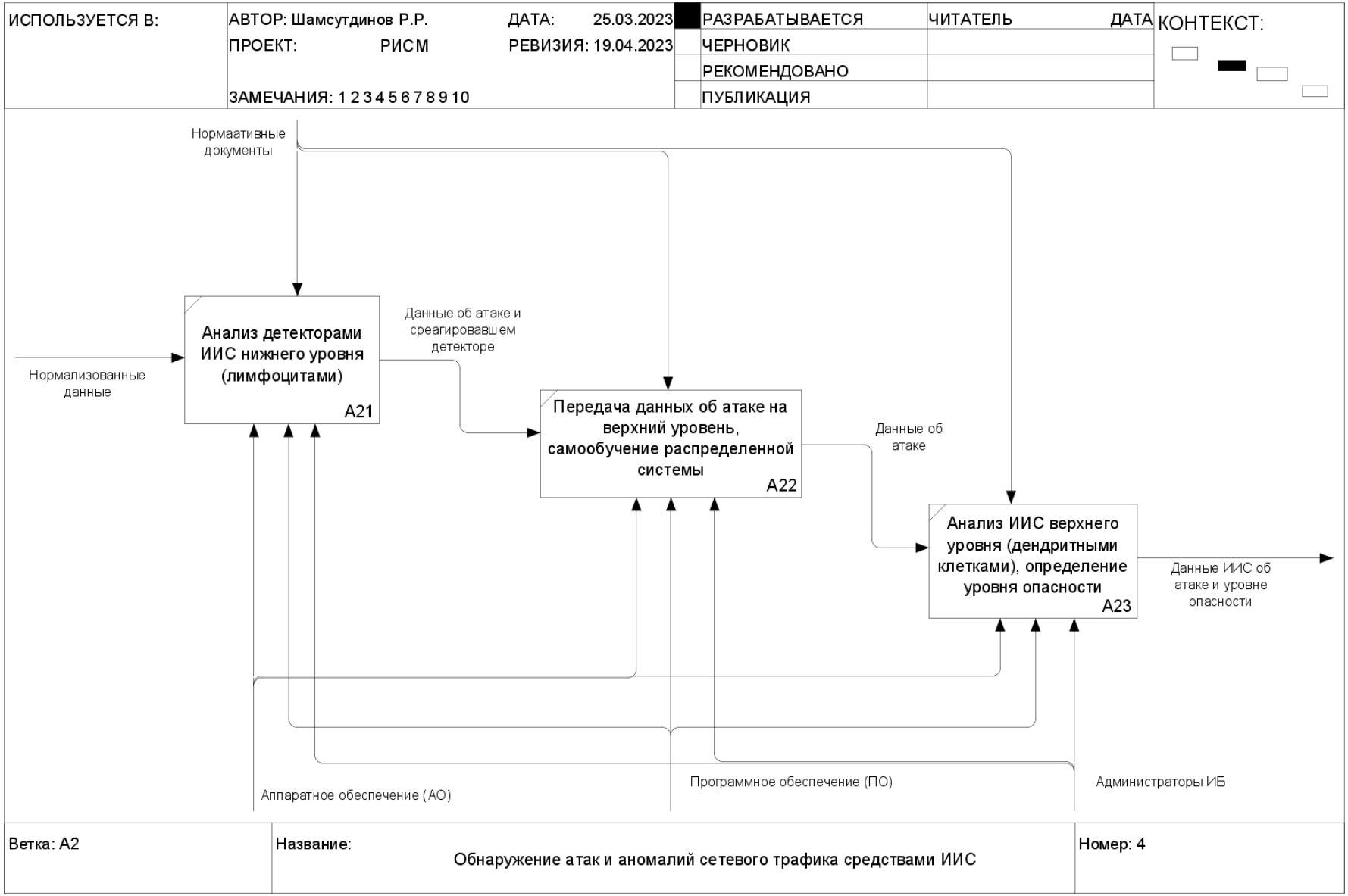


Рисунок 2.1 (лист 4 из 6)

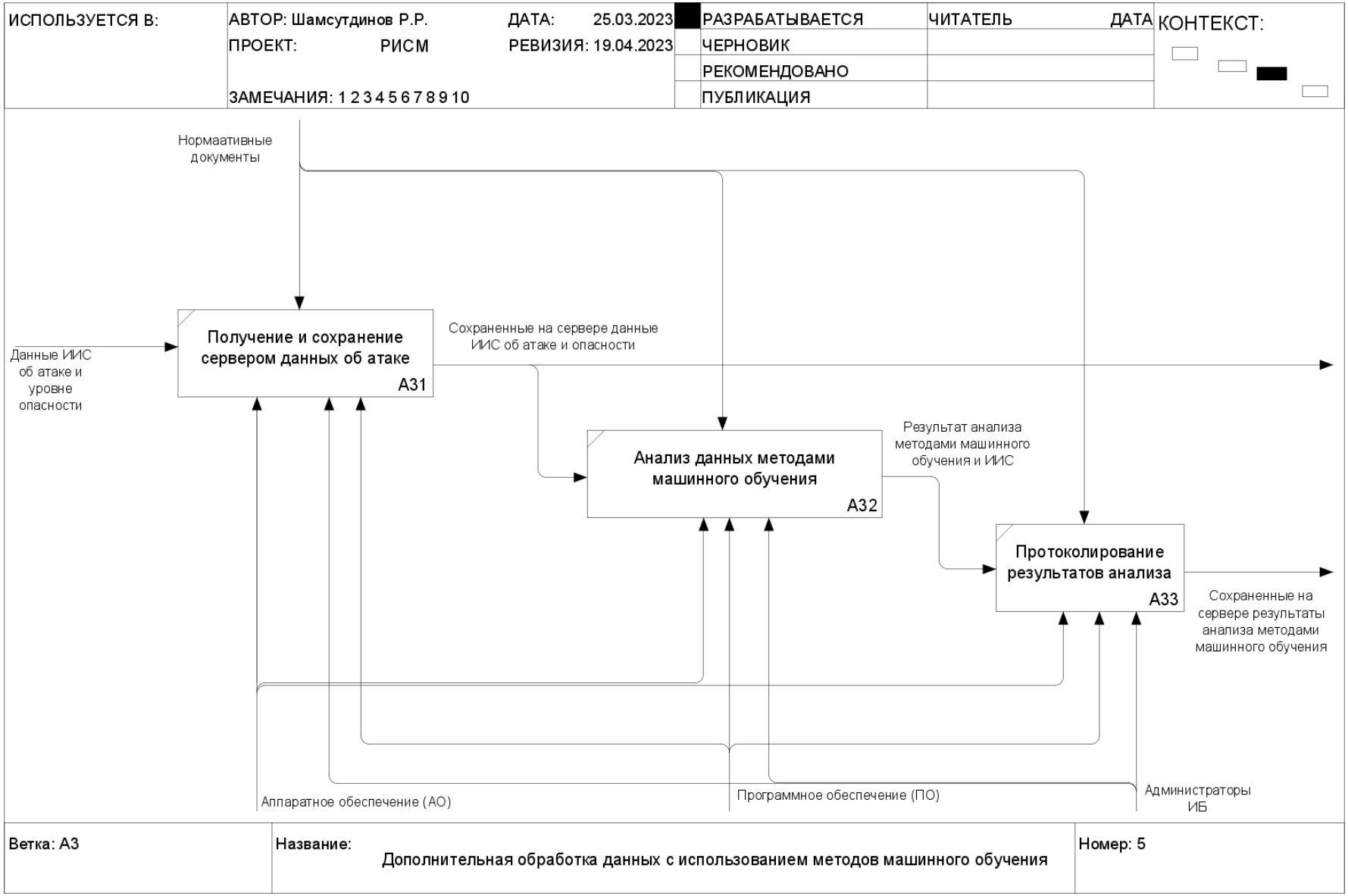


Рисунок 2.1 (лист 5 из 6)

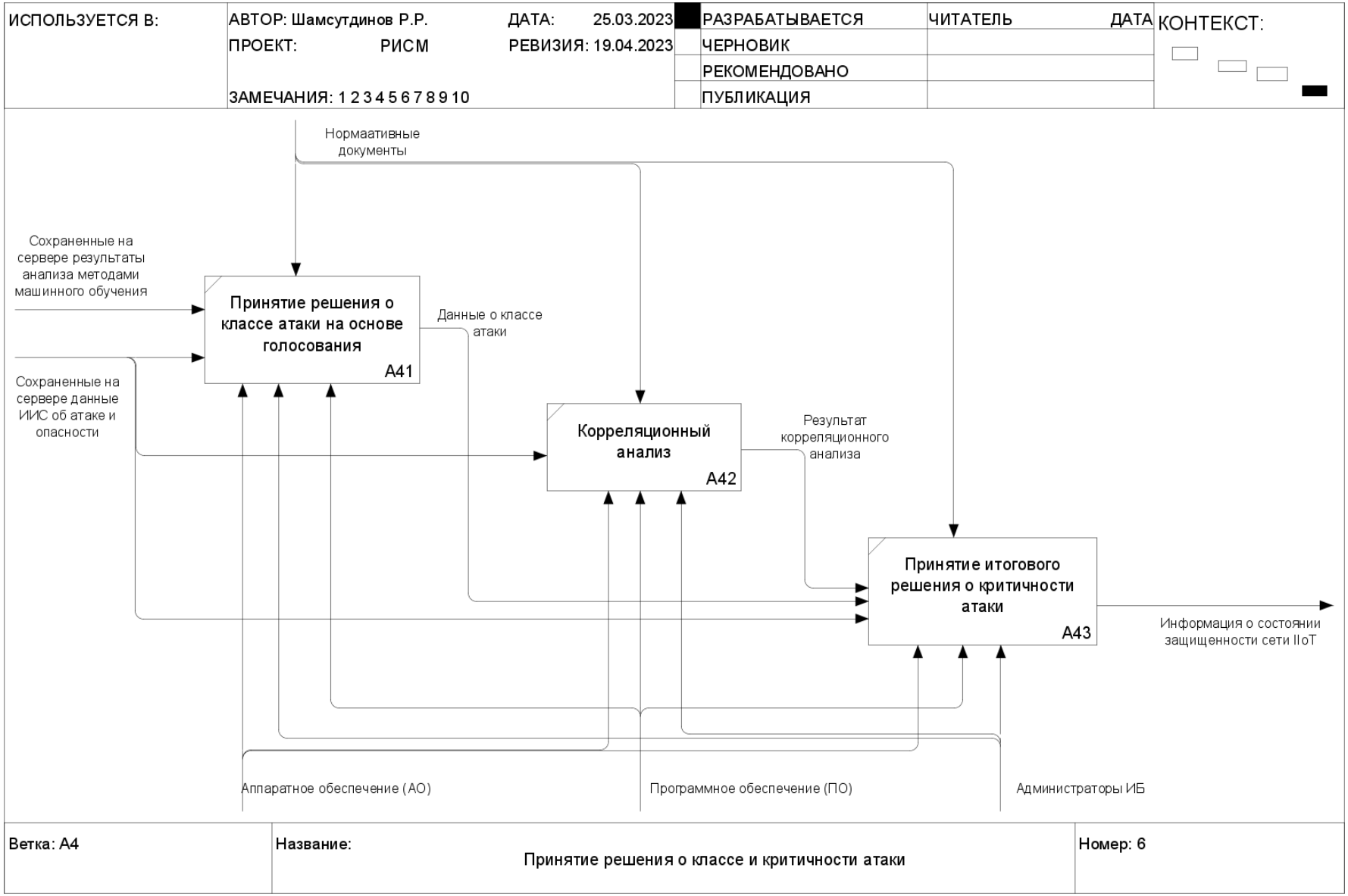


Рисунок 2.1 (лист 6 из 6)

Двухуровневая ИИС также содержит агенты верхнего (второго) уровня, реализующие вычисления на основе принципов теории опасности в виде алгоритмов дендритных клеток (ДК), агрегирующие данные об атаках от подконтрольных агентов нижнего уровня, анализирующие уровень опасности.

UML-диаграмма классов двухуровневой ИИС представлена на Рисунке 2.2. Основные классы модели ИИС приведены в Таблице 2.1.

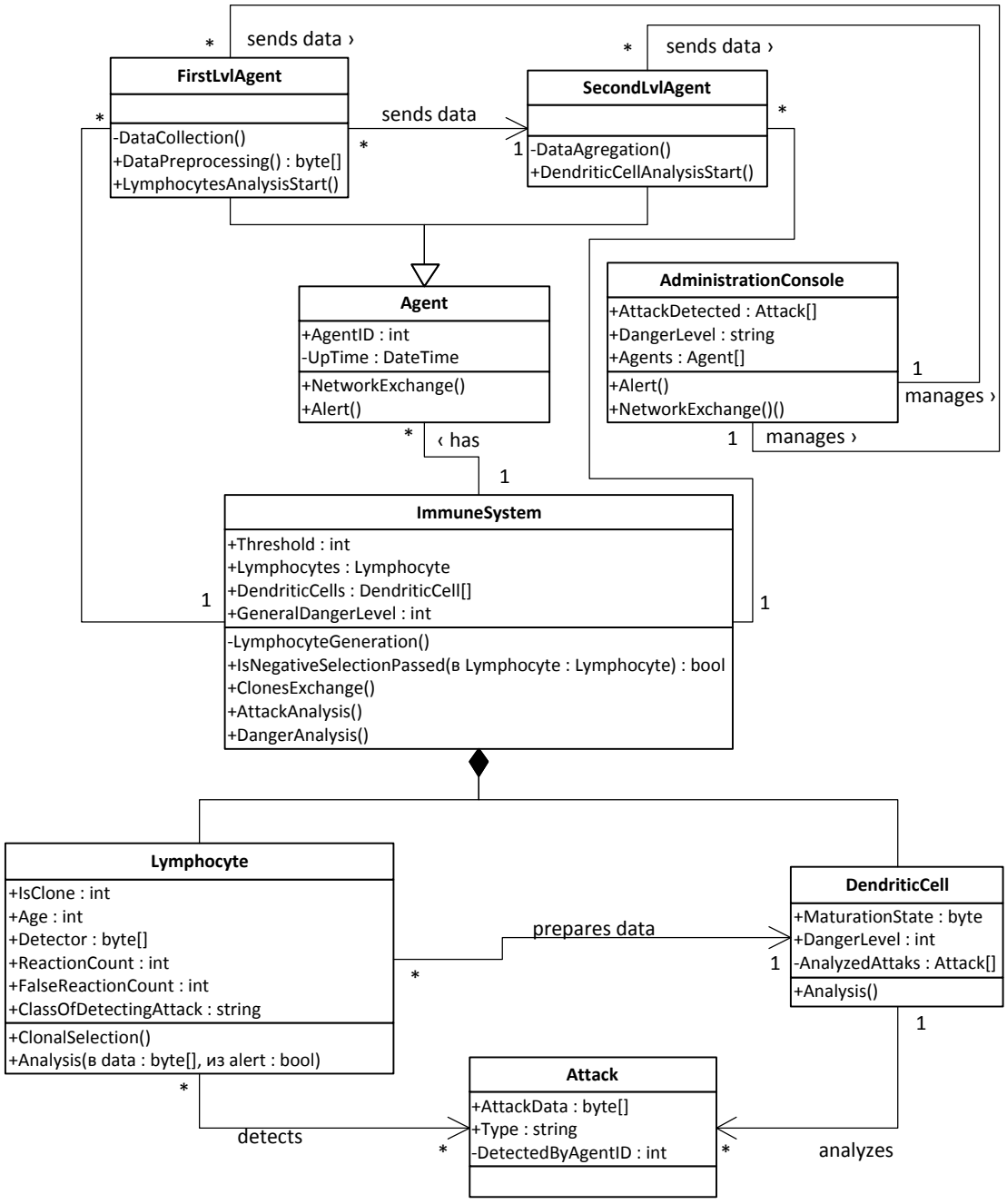


Рисунок 2.2 – Диаграмма классов ИИС

Таблица 2.1 – Основные классы модели ИИС

Класс	Функционал
ImmuneSystem	основной класс ИИС, реализующий совместно с дочерними классами алгоритмы ИИС
Lymphocyte	дочерний класс от ImmuneSystem. Реализует выявление атак и аномалий, их первичную классификацию, самообучение.
DendriticCell	дочерний класс от ImmuneSystem. Реализует анализ уровня опасности посредством реализации алгоритма дендритных клеток.
Attack	класс, представляющий собой совокупность данных о выявленной атаке
Agent	общий родительский класс агентов, реализующий идентификационные атрибуты
FirstLvlAgent	реализует сбор данных и запускает функционирование алгоритмов класса Lymphocyte.
SecondLvlAgent	агрегирует данные, получаемые от класса FirstLvlAgent, инициирует алгоритмы дендритных клеток класса DendriticCell
AdministrationConsole	класс, реализующий алгоритмы консоли администрирования, в том числе прием данных и оповещение об атаках

Таким образом, рассмотренная двухуровневая ИИС, согласно Рисунку 2.2 и Таблице 2.1, содержит два вида агентов: первого и второго уровней, использующие методы классов «лимфоцит» и «дендритная клетка» соответственно. Агенты первого уровня, используя метод анализа класса «лимфоцит», подготавливают данные для агентов второго уровня, который реализует метод анализа класса дендритных клеток.

2.2 Сбор и предварительная обработка данных о сетевом трафике промышленного Интернета вещей

В данной работе была использована следующая процедура определения информативных (значимых) параметров датасета. Лингвистические данные были кодированы целыми неотрицательными числами, получился диапазон значений [0; 65]; десятичные дроби, округленные до сотых, были умножены на 100; флаговые значения оставлены без изменений; остальные целочисленные

неотрицательные значения – распределены в диапазоне [0; 255] следующим образом. Для каждого параметра было выбрано пороговое значение, близкое к некоторому максимальному. Если исходное значение строго равно нулю, то «сжатое» значение также равно нулю. Если исходное значение больше нуля и не превышает порогового, то сжатие осуществляется равномерно так, чтобы сжатое значение лежало в диапазоне [1; 254]. Если исходное значение превышает пороговое, то сжатое равно 255. То есть:

$$y_i = \begin{cases} 255, & \text{если } x_i > P_i \\ 1 + \left\lfloor x_i : \left(\frac{P_i}{253}\right) \right\rfloor, & \text{если } 0 < x_i \leq P_i \\ 0, & \text{если } x_i = 0 \end{cases}, \quad (2.1)$$

где y_i – значение параметра после сжатия; x_i – значение параметра до сжатия; P_i – пороговое значение параметра.

В итоге, значение каждого параметра после нормализации представляет собой целое неотрицательное число в диапазоне значений одного байта.

Таким образом, исходный набор данных был нормализован, разделен на данные об атаках (*A)* и о нормальном состоянии (*N)*. Для каждой строки A_i была найдена максимально похожая строка N_j в датасете, в качестве меры близости использовалось расстояние Хэмминга. Совпадающие параметры для каждой такой пары строк были отмечены. После чего был выполнен расчёт частоты совпадений по каждому параметру, ранжирование по наименьшей частоте совпадений.

На следующем этапе предполагается обучение используемого классификатора (СОА) на основе выбранного количества ранжированных параметров и оценка точности классификации. При ее недостаточности требуется увеличение числа параметров. Если точность классификации достаточна, стоит уменьшить количество параметров и повторить эксперимент для определения рационального количества анализируемых параметров.

Блок-схема разработанного алгоритма нормализации данных датасета приведена в Приложении Б (Рисунки Б.1-Б.2), алгоритма нормализации данных реально перехваченного или полученного трафика – Рисунок Б.3.

Таким образом, обработка трафика является двухуровневой: в первую очередь извлекаются выбранные параметры из файлов с расширением pcap (.pcap-файлов), NetFlow или других источников; если требуется, вычисляются дополнительные параметры. Затем параметры приводятся к анализируемому формату представления. Данные сетевого трафика должны поступать анализатору непрерывно, аналогично поступлению данных в SIEM-системы.

Преимуществами предложенного алгоритма являются его простота: не требуется построения сложных классификаторов, а также возможность более гибко регулировать границы распределения нормализованных значений для параметров с большой величиной разброса.

Рассмотрим упрощенный пример. Допустим, что большинство строк значений определенного параметра принимают значения в диапазоне [0; 200], а для некоторых строк его значения равны 5 000, 8 000, 10 000 и т.д., и пусть диапазон значений будет сжат в целых числах равномерно в 100 раз, то есть с [0; 10 000] до [0; 100]. Тогда весь диапазон значений [0; 200] превратится в [0; 2], и будут потеряны значимые данные, так как разницы между исходными значениями 156 и 199 после такого уменьшения диапазона и округления уже не будет.

Кроме того, значение, равное единице, при значительном уменьшении используемого диапазона почти наверняка будет приведено к нулю, а разница между нулевым значением и ненулевым может быть существенной. Поэтому в рассматриваемом примере выбирается пороговое значение 200, исходное нулевое значение сохраняется нулевым, все значения больше 200 приняты равными 100, а диапазон [1; 200] равномерно сжат до диапазона [1; 99].

Стоит отметить, что применение подсчёта частоты совпадений параметров при определении их значимости во многом аналогичен метрике Хэмминга, что позволяет сохранить эффективность разделения множеств атак и нормального состояния с использованием меры Хэмминга, часто используемой в ИИС.

Данный алгоритм не предполагает какой-либо балансировки данных, содержащихся в датасете. Для многих интеллектуальных классификаторов необходимо наличие определенного количества образцов атак каждого класса для эффективного обучения, однако большинство датасетов содержит в том числе атаки, для которых представлено недостаточное количество образцов. Данная проблема решается в [80], где авторами предлагается применение алгоритма генеративных состязательных сетей для дополнения малочисленных атак сгенерированными образцами.

В рамках планируемых в работе вычислительных экспериментов по оценке эффективности использования ИИС для обнаружения сетевых атак и аномалий будет использоваться несбалансированный датасет для дополнительного определения возможности ИИС выявлять атаки, представленные в малом количестве.

Таким образом, рассмотрен алгоритм выбора и нормализации данных сетевого трафика на основе анализа параметров сетевых соединений, используемых в датасетах, предназначенных для построения и тестирования систем обнаружения атак, основанный на приведении исходных значений параметров к заданному целочисленному диапазону, разделении датасета на подмножества данных об атаках и данных о нормальном сетевом взаимодействии, определении частоты совпадений параметров между подмножествами, ранжировании по наименьшему проценту совпадений (наибольшей информативности), проведении вычислительных экспериментов по обнаружению атак с использованием различного количества наиболее информативных параметров.

2.3 Алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика на основе адаптивных механизмов искусственных иммунных систем

Искусственная иммунная система (ИИС) имитирует работу естественной иммунной системы человека, предназначенной для его защиты от внешних и внутренних, известных и неизвестных угроз. При наиболее классическом варианте построения ИИС реализуется классификация по принципу «свой/чужой». В таком случае строится множество детекторов, каждый из которых содержит некоторую вектор-строку, являющуюся реальным или предполагаемым эталоном атаки.

Такая строка может составляться непосредственно из параметров сетевых соединений, соответствующих атакам или аномалиям, или генерироваться случайным образом, но при обеспечении ее уникальности и гарантии ее существенного отличия от данных, соответствующих нормальному сетевому взаимодействию.

ИИС анализирует данные в формате некоторой последовательности значений. Это может быть последовательность определенных событий или набор параметров, характеризующих одно событие, но в любом случае данные должны быть представлены в виде вектора-строки, который также определяется точкой в пространстве параметров. Анализ заключается в определении схожести векторов анализируемых данных и эталонных векторов детекторов «чужого». Если векторы достаточно близки, считается, что анализируемый экземпляр данных соответствует некоторой аномалии.

ИИС является адаптивной системой, если в ней реализованы механизмы клональной селекции, обеспечивающие клонирование полезных (обнаруживших атаку или аномалию) детекторов с искажением эталонного вектора-строки для лучшего обнаружения подобных угроз, что позволяет ИИС постоянно самообучаться. ИИС также легко адаптировать для построения в качестве

распределённой системы взаимодействующих независимых децентрализованных узлов.

Как правило, пространство параметров является довольно большим и генерация детекторов, покрывающих все пространство, кроме точек, соответствующих нормальному состоянию, требует значительных вычислительных ресурсов.

Поэтому довольно часто применяются механизмы обновления состава детекторов и формирования детекторов памяти. Детекторы, которые за определенное время не обнаружили какой-либо атаки или аномалии, уничтожаются, что аналогично естественной смерти иммунных клеток в естественной иммунной системе. В ином случае, если детектор обнаружил атаку или аномалию, он превращается в детектор памяти и срок его жизни значительно увеличивается, подобно клеткам памяти естественной иммунной системы, которые обеспечивают быструю реакцию нейтрализации уже известного патогена.

При построении ИИС, функционирующей по принципу «опасно/безопасно» детекторы имитируют работу так называемых дендритных клеток, анализирующих наличие или отсутствие сигналов опасности, безопасности и сигналов о наличии патогенов (PAMP – Pathogen Associated Molecular Pattern). Результатом анализа является вывод об опасности или безопасности выявленного патогена и активации процессов его нейтрализации или выработки к нему толерантности соответственно.

Требования, предъявляемые к системе обнаружения атак:

– адаптивность – возможность постоянного обновления базы детекторов;

– относительно невысокая сложность реализации, низкие требования к памяти, высокая скорость обнаружения;

– высокая эффективность обнаружения атаки (как из числа известных, так и ранее неизвестных).

Стоит отметить, неизвестные атаки, которые можно обнаружить на основе данных сетевого трафика, выявляются на основе обнаружения аномалий. Если параметры сетевого трафика в определенный момент времени отличаются от некоторых шаблонов нормального взаимодействия, значит, возникла аномалия, которая может быть вызвана единичным безвредным явлением, или же атакой, не известной системе.

Когда в работе идет речь о неизвестных атаках, имеются в виду атаки, не известные системе. Потому что любая атака, даже сверхновая как минимум известна её автору, поэтому абсолютно никому не известные атаки представить сложно, вероятно, их и не существует. Речь идет об известности атаки конкретной системе обнаружения.

Нередки случаи, когда осуществляется новая атака, которая наносит ущерб, становится известной в кругах информационной безопасности, но для её обнаружения ещё нет сигнатур (атака нулевого дня). И системы защиты, которые не имеют потенциала к обнаружению неизвестных для себя атак, ее не выявляют. Поэтому важно, чтобы система обладала потенциалом к обнаружению неизвестных для себя атак, была адаптивной.

Для оценки эффективности системы использовались следующие метрики [68, 141]:

– False Negatives (FN) – количество образов атак, определенных как норма (ошибки второго рода);

– False Positives (FP) – количество образов нормальной активности, определенных как атаки (ошибки первого рода);

– True Negatives (TN) – количество верно определенных образов нормальной активности;

– True Positives (TP) – количество верно выявленных атак;

– False Negative Rate (FNR) – уровень ошибок второго рода:

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{TP+FN}}; \quad (2.2)$$

– False Positive Rate (FPR) – уровень ошибок первого рода:

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN+FP}}; \quad (2.3)$$

– True Negative Rate (TNR) – доля верно определенных образов нормальной активности:

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{TN+FP}}; \quad (2.4)$$

– Recall (полнота, True Positive Rate, TRP, также обозначается как Sensitivity – чувствительность) – доля верно выявленных атак среди всех атак:

$$\mathrm{Recall} = \mathrm{TRP} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}; \quad (2.5)$$

– Precision (точность) – доля верно выявленных атак среди всех образов, определенных как атаки:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}; \quad (2.6)$$

– Accuracy – доля верно классифицированных образов среди всех образов:

$$\mathrm{Accuracy} = \frac{\mathrm{TP+TN}}{\mathrm{TP+FP+TN+FN}}; \quad (2.7)$$

– F_1 score – среднее гармоническое точности (Precision) и полноты (Recall):

$$F_1 score = \frac{2 \times Precision \times Recall}{Precision + Recall}. \quad (2.8)$$

Под эффективностью системы обнаружения атак и аномалий здесь и далее в работе понимается совокупность значений показателей эффективности: Recall, Precision, Accuracy, F_1 score. Можно говорить об увеличении эффективности, если увеличиваются значения всех этих показателей, или если значения одних показателей увеличиваются, а оставшихся – не уменьшаются. Эффективность снижается, если значение хотя бы одного из этих показателей снижается.

В целом, алгоритм функционирования разработанной искусственной иммунной системы (ИИС) представлен Рисунком 2.3.

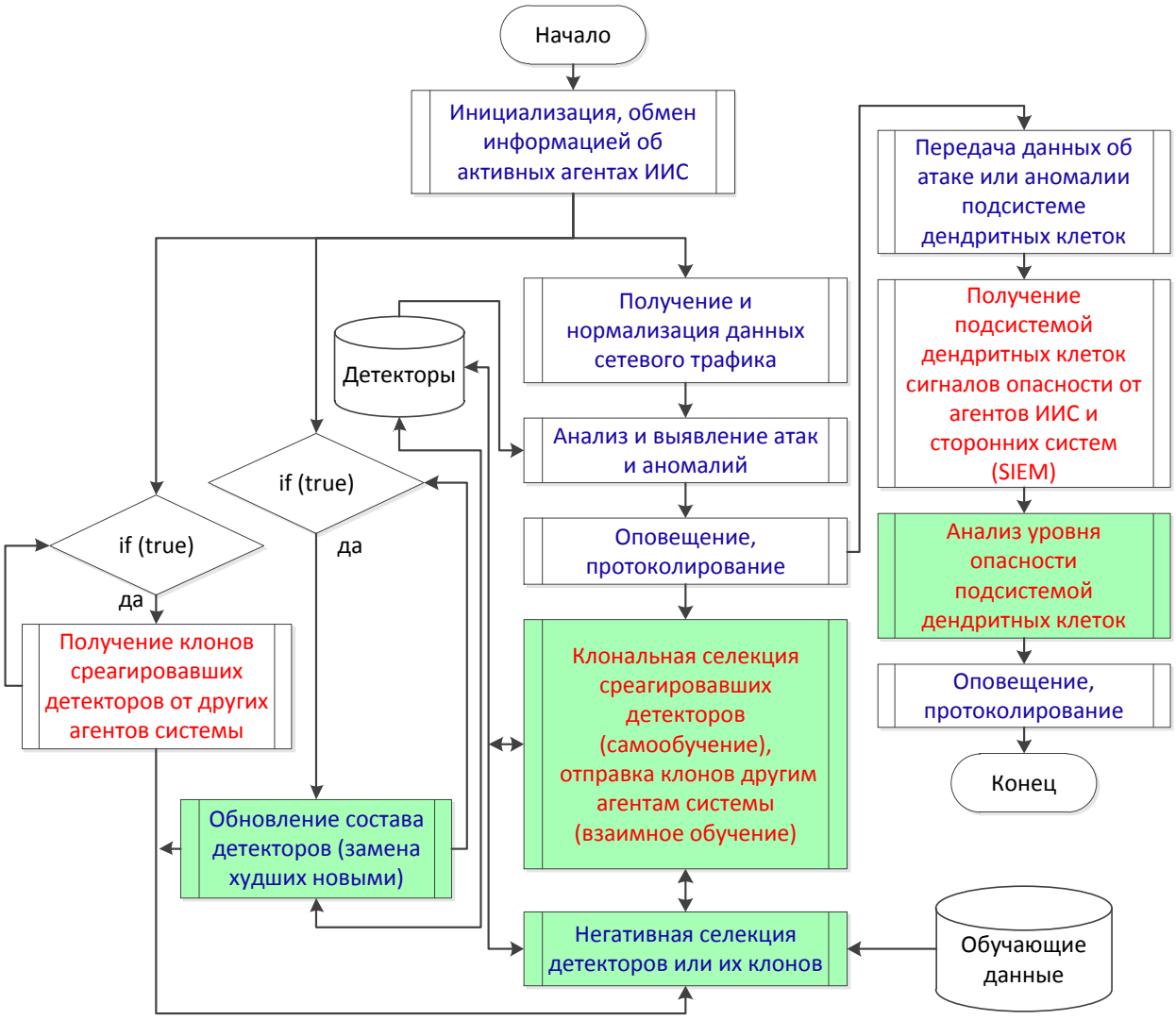


Рисунок 2.3 – Алгоритм функционирования разработанной ИИС

На Рисунке 2.3 шрифтом тёмно-синего цвета выделены подпроцессы, построенные на основе существующих алгоритмов, шрифтом красного цвета – новые и доработанные алгоритмы.

В целом, новым является:

– объединение различных алгоритмов ИИС, выделенных на Рисунке 2.3 зелёной заливкой, в единой системе (подпроцессы обновления состава детекторов, негативной селекции детекторов или их клонов, клональной селекции среагировавших детекторов и отправки клонов другим агентам системы, анализа уровня опасности подсистемой дендритных клеток, отображенные на рисунке 2.3);

– модификация алгоритма клональной селекции для функционирования в многоагентной распределенной архитектуре таким образом, что каждый агент, проводящий анализ и выявление атак и аномалий, на основе результатов анализа не только создает видоизменённые клоны лучших детекторов, заменяющих собой худшие детекторы (самообучается), но и передаёт полученные клоны всем другим агентам нижнего уровня, обеспечивая взаимное обучение агентами друг друга (подпроцессы получения клонов среагировавших детекторов от других агентов системы, клональной селекции среагировавших детекторов и отправки клонов другим агентам системы, отображенные на рисунке 2.3);

– разработанные механизмы взаимодействия подсистемы дендритных клеток с подсистемами корреляционного анализа – SIEM-системами (подпроцесс получения подсистемой дендритных клеток сигналов опасности от агентов ИИС и сторонних систем) рассматриваются в данном разделе, а также в разделе 3.2;

– модификация алгоритма дендритных клеток, исключающая анализ сигналов безопасности, блокирующая уязвимость, связанную с навязыванием ложных сигналов безопасности для обеспечения видимости легитимного сетевого взаимодействия во время атаки (подпроцесс анализа уровня опасности подсистемой дендритных клеток, отображенный на рисунке 2.3).

Перед обнаружением атак ИИС следует в первую очередь ее обучить, соответствующий алгоритм обучения представлен блок-схемой на Рисунке 2.4.

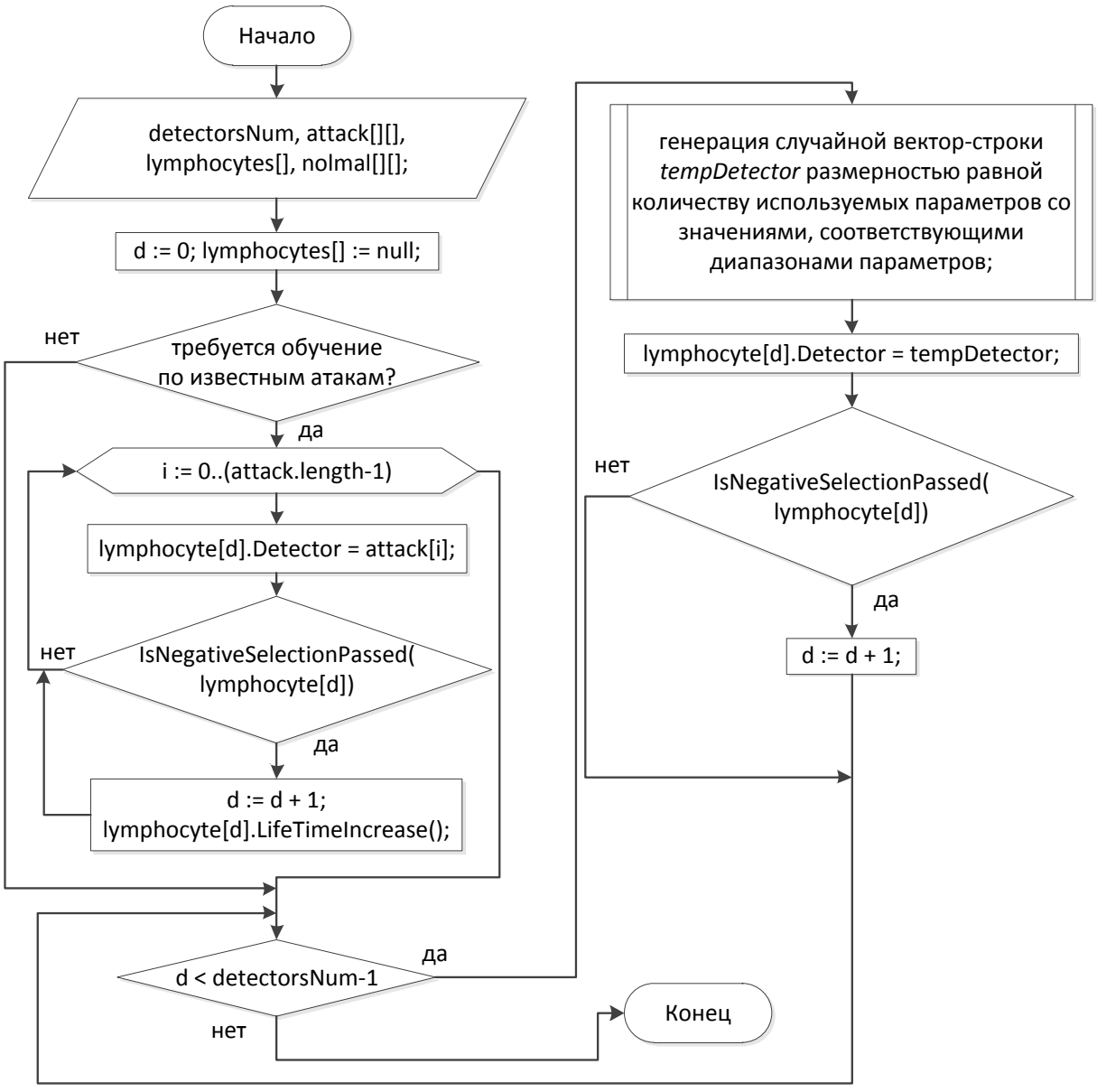


Рисунок 2.4 – Блок-схема алгоритма обучения ИИС

В первую очередь определяется, будет ли система обучена только на основе данных нормального трафика и в итоге будет только выявлять атаки, воспринимая их все как неизвестные, или же система будет обучена в том числе на данных об атаках и будет классифицировать данные трафика как соответствующие нормальному взаимодействию, определенному классу атаки, или не

соответствующие ни одному из этих классов, то есть представляющие собой проявление аномалии или неизвестной атаки. В первом случае, каждый пример данных об атаках становится потенциальным детектором, который далее в целях его удаленности от векторов нормальных данных и от других детекторов подвергается негативной селекции, блок-схема алгоритма реализации которого представлена на Рисунке 2.5.

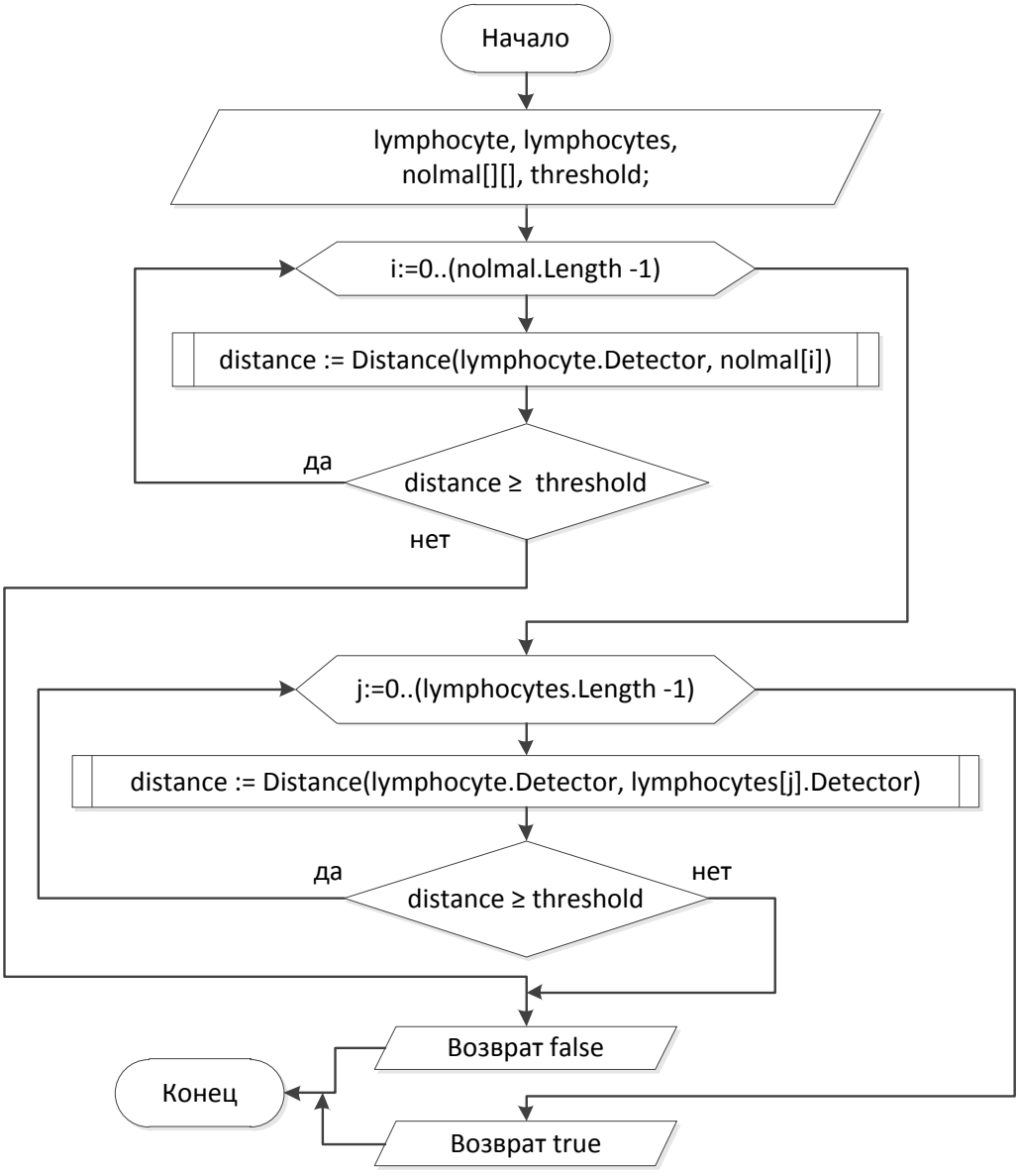


Рисунок 2.5 – Блок-схема алгоритма негативной селекции

Программно негативная селекция реализуется функцией *IsNegativeSelectionPassed(lymphocyte),* возвращающей булево значение: пройдена

она успешно или нет. Если пройдена, то детектор сохраняется, время его существования увеличивается функцией *Lymphocyte.LifeTimeIncrease()*; поскольку он точно соответствует какой-либо атаке, целесообразно его хранить длительно. После завершения обучения на основе набора данных об атаках или же в случае отказа от его использования, множество лимфоцитов дополняется до заданной размерности экземплярами со сгенерированной случайным образом строкой-детектором. Такие лимфоциты также подвергаются негативной селекции.

В блок-схеме алгоритма на Рисунке 2.5 используется функция *Distance(vector1, vector2)*, с помощью которой находится расстояние между двумя вектор-строками, служащими ее аргументами. В качестве этой функции допускается использование любой меры расстояния между векторами, в данной работе использовано расстояние Хэмминга, *affinityThreshold* – пороговое значение близости (аффинности) векторов, при превышении которого векторы определяются как близкие относительно друг друга; *threshold* – обратная величина (порог удаленности), достижение которого свидетельствует о значительном расстоянии между векторами, вычисляется как разность размерности вектора n и порога аффинности a.

Первичный анализ входных данных и выявление атак и аномалий лимфоцитами осуществляется в целом с помощью алгоритма, блок-схема которого представлена на Рисунке 2.6. В соответствии с этим алгоритмом, в первую очередь осуществляется перехват и нормализация данных о текущем сетевом взаимодействии.

Затем детектором каждого существующего лимфоцита выполняется анализ: если входные данные соответствуют аномалии или атаке, они передаются дендритным клеткам для анализа опасности, также для среагировавших лимфоцитов выполняется процедура клональной селекции *Lymphocyte.ClonalSelection()*.

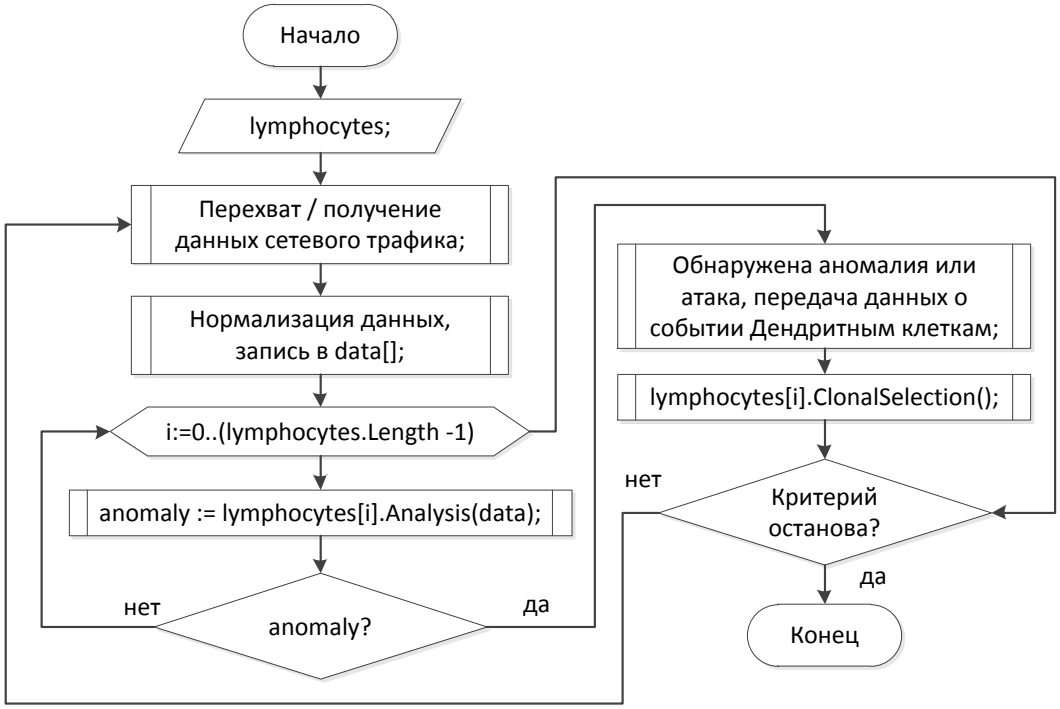


Рисунок 2.6 – Укрупненная блок-схема алгоритма анализа входных данных

Анализ лимфоцитом выполняется с использованием функции *Lymphocyte.Analysis()*, возвращающей булево значение о том, обнаружена лимфоцитом атака (аномалия) или нет. Блок-схема алгоритма реализации данной функции представлена на Рисунке 2.7.

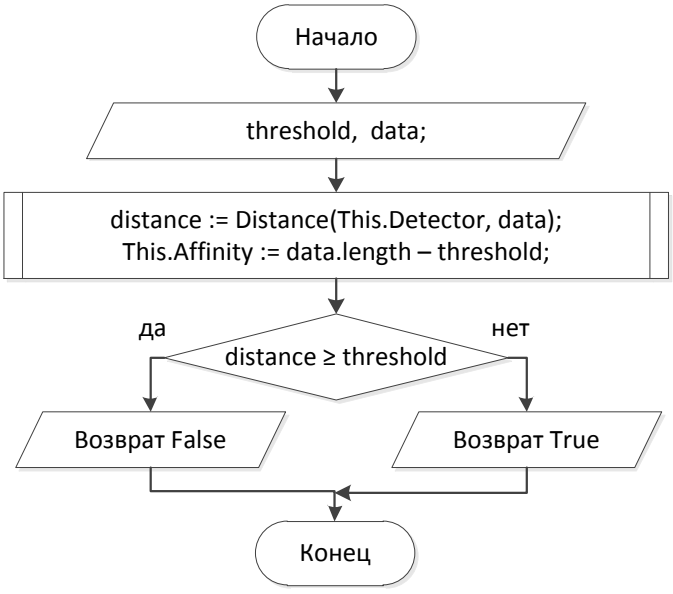


Рисунок 2.7 – Блок-схема алгоритма определения реакции лимфоцита

Реализация функции *Lymphocyte.Analysis()* основана на использовании рассмотренной ранее функции *Distance*. Если расстояние между векторами достигает порогового, считается, что детекторы-строки не реагируют и обнаружена легитимная активность, иначе – аномальная.

Процедура клональной селекции, представленная блок-схемой алгоритма на Рисунке 2.8, заключается в создании видоизмененных (мутировавших) клонов исходного детектора в количестве, прямо пропорциональном его аффинности (близости) к шаблону соответствующей аномалии или атаки.

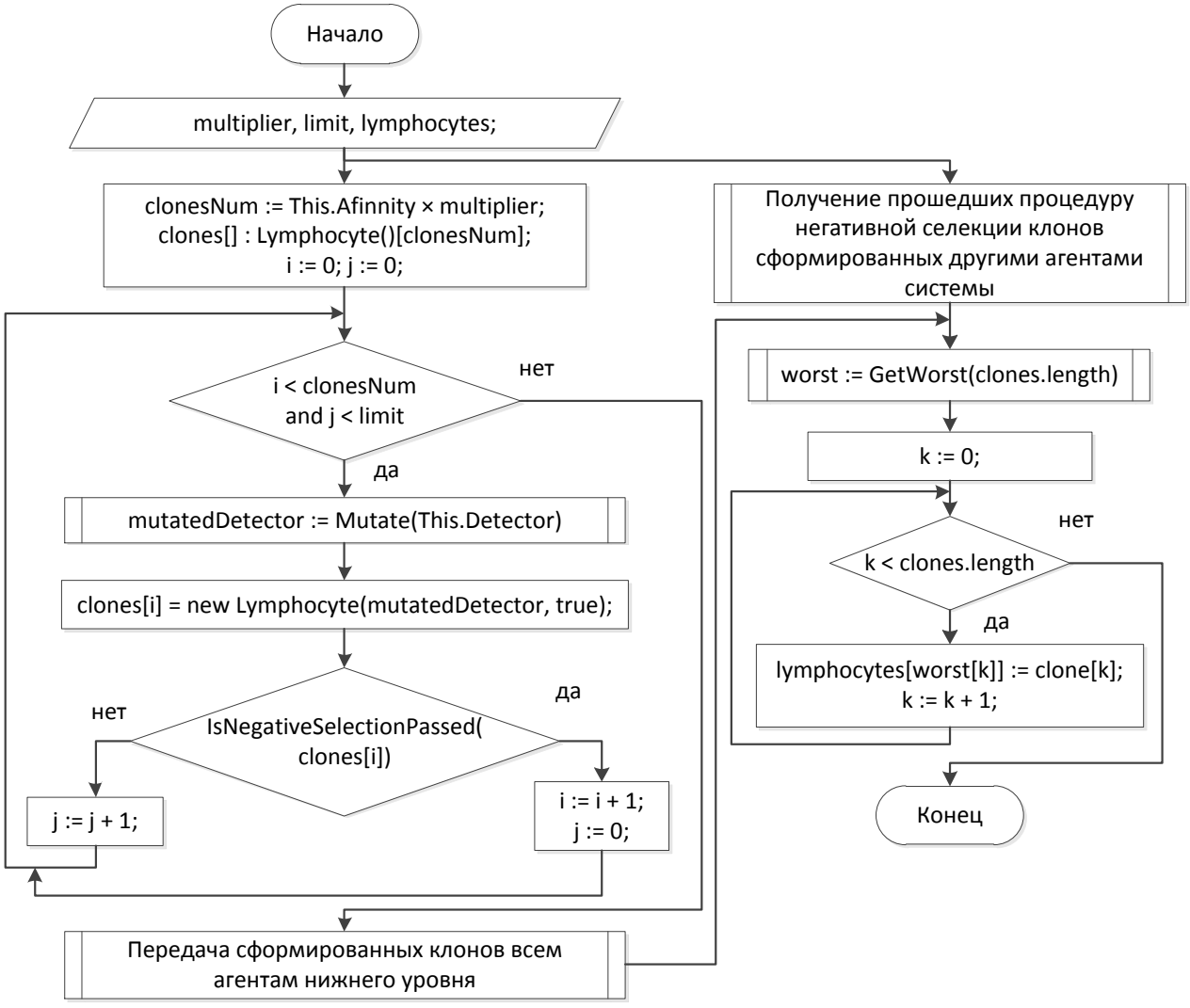


Рисунок 2.8 – Блок-схема алгоритма реализации функции *Lymphocyte.ClonalSelection()*

Мутация осуществляется с помощью функций *Mutate(Lymphocyte)*, которая выбирает случайный номер параметра (компоненты вектора) и записывает в качестве его значения случайное число, соответствующее диапазону значений данного параметра. Каждый создаваемый клон также подвергается негативной селекции; если клон ее не проходит, создается новый, пока не будет достигнуто необходимое количество клонов или пока не будет выполнено определенное количество попыток, заданное константой *limit*.

Функция *GetWorst()* возвращает заданной размерности массив номеров наихудших лимфоцитов, имеющих наибольшее значение возраста и наименьшее количество выявленных атак или аномалий.

Полученные клоны отправляются всем другим агентам системы, где снова проходят негативную селекцию и пополняют собой множество детекторов других агентов. Данное нововведение не является каким-либо сложным алгоритмом для представления блок-схемы его функционирования, но получен существенный эффект от его реализации, подробнее рассматриваемый в вычислительных экспериментах в разделе 4.3.

После выявления аномалий или потенциальных атак, информация о них передается подсистеме дендритных клеток (ДК). Классический алгоритм ДК подразумевает обеспечение толерантности к обнаруженной атаке или аномалии и определение ее как безвредной, если уровень сигналов безопасности, связанных с данной атакой или аномалией, превышает уровень сигнала опасности. Это порождает значительную уязвимость: злоумышленник может имитировать сигналы безопасности, проводя атаку, тогда система безопасности не среагирует на атаку не только в данном случае, но и во всех последующих, так как выработает толерантность к ней. Поэтому указанный алгоритм требует доработки.

Алгоритм ДК был переработан следующим образом, представленным на Рисунке 2.9. В качестве сигналов опасности и безопасности можно определить различные события ИБ; в случае возможности установки некоторого агента на само устройство IoT или возможности его настройки на передачу данных о своем

состоянии, безопасность или уровень опасности можно оценивать на основе данных о состоянии контролируемых устройств.

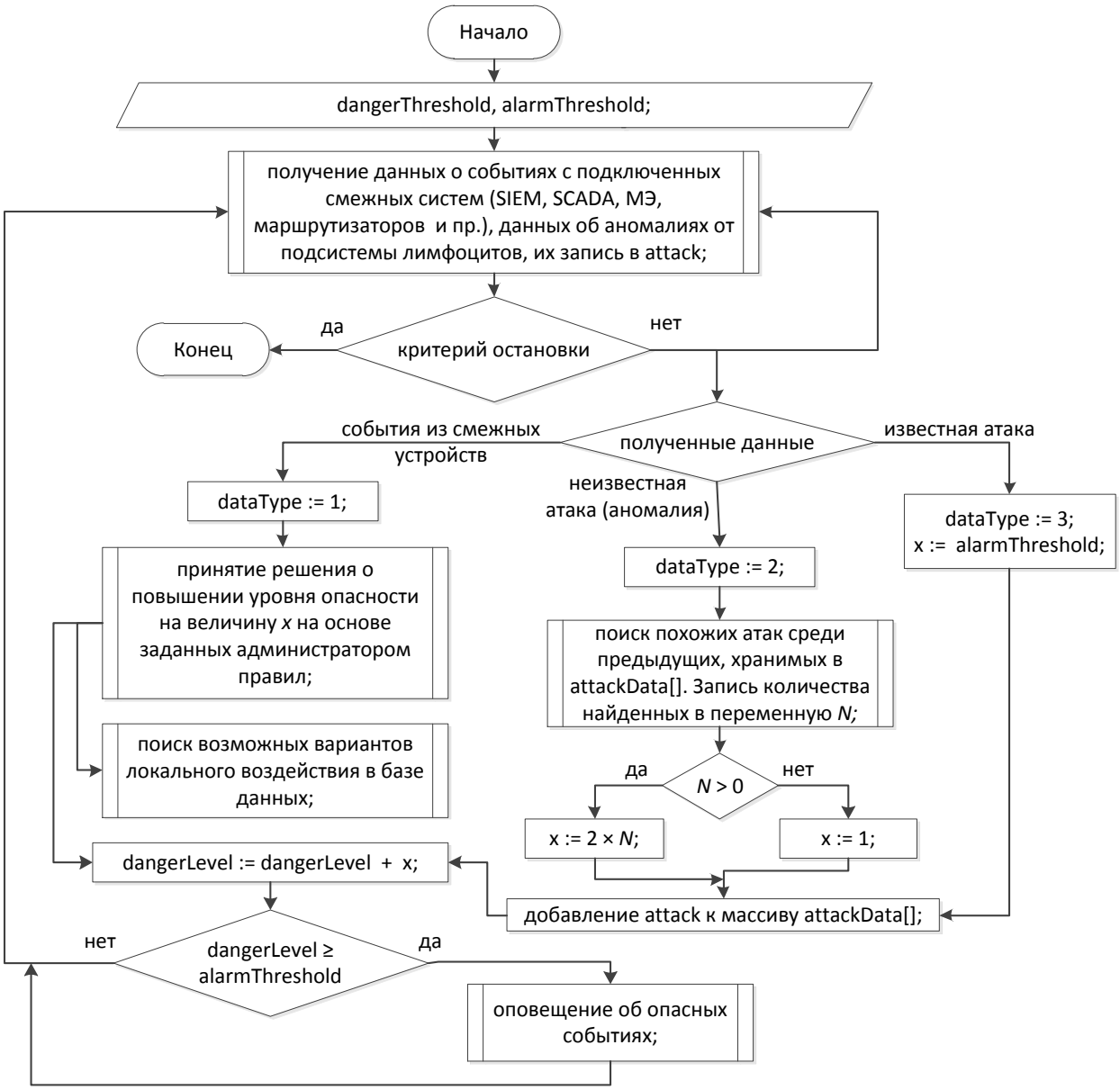


Рисунок 2.9 – Блок-схема алгоритма реализации функции *DendriticCell.Analysis()*

Также дополнительными источниками подобных сигналов могут быть, например, данные от межсетевого экрана (МЭ) о количестве заблокированных соединений за единицу времени: чем выше этот показатель, тем выше уровень опасности, SIEM-системы, SCADA-системы. Так, если некоторый датчик зафиксировал отклонение управляемых координат ТП от планируемых и

системой автоматизации должно быть инициировано управляющее воздействие для компенсации отклонения, но оно не было выполнено, то повышается уровень опасности.

Если контролируемая система функционирует в штатном режиме, уровень опасности равен нулю, и вдруг возникла неизвестная аномалия, информация о ней сразу попадает в ДК, при этом незначительно (на единицу) повышается уровень опасности, определяемый этой клеткой, но он по-прежнему остается близок к нулю и не вызывает каких-либо реакций. Если возникает еще одна аномалия, она дополнительно повышает уровень опасности еще на единицу, но если обнаруженные аномалии сильно подобны, уровень опасности повышается сильнее:

$$\Delta DangerLevel = 2 \times N, \tag{2.9}$$

где $\Delta DangerLevel$ – прирост уровня опасности, N – количество подобных атак или аномалий, выявленных ранее.

При большом потоке идентичных аномалий, резко возрастает уровень опасности, аномалии классифицируются как неизвестная атака, инициируется оповещение, возможно формирование каких-либо управляющих воздействий.

Можно обеспечить возможность задания вручную того условия, что именно и в какой степени считать опасным; к примеру, потенциально опасной можно считать заблокированную МЭ попытку подключения по SSH к конкретному устройству, особенно по внутренней сети.

Если лимфоциты обучены, в том числе на основе данных об атаках, и событие ИБ выявлено одним из таких лимфоцитов или его клонов, это событие определяется не как аномалия, а как атака, и резко повышается уровень опасности. С течением времени при отсутствии угрожающих событий уровень опасности постепенно снижается. Данный алгоритм отличается адаптацией к работе с подсистемой лимфоцитарного анализа, различием обработки данных об известных атаках и о потенциально возможных атаках, возможностью интеграции

с различными сторонними системами мониторинга, отсутствием анализа сигналов безопасности.

Была построена двухуровневая ИИС, которая была обучена и протестирована с использованием набора данных NSL-KDD в двух режимах:

1) система обучалась на основе данных и о нормальной активности, и об атаках;

2) система обучалась только на основе данных о нормальной активности, и все атаки были для нее неизвестны.

NSL-KDD содержит 22 вида атак, представленных в Таблице 1.4, объединенных в 4 группы:

– Denial of Service (DoS) – отказ в обслуживании, представляет собой нарушение доступности системы, являющейся объектом атаки, посредством генерации большого количества запросов или обращений к системе;

– User to Root (U2R) – превышение полномочий, представляет собой несанкционированное получение прав более привилегированного пользователя при уже имеющемся доступе в качестве рядового пользователя;

– Remote to Local (R2L) – получение удаленного доступа к системе извне;

– Probe – сканирование, представляет собой сбор информации о вычислительной сети с целью дальнейшего обхода систем обеспечения её безопасности.

Выбранный датасет NSL-KDD был нормализован рассмотренным в разд. 2.2 способом, с разделением данных о нормальной активности и данных об атаках. Подсистема лимфоцитов, содержащая детекторы, реализующие анализ по принципу «свой/чужой», предварительно обучалась на 50% данных о нормальной активности и 50% данных об атаках для возможности их корректной классификации при сохранении возможности обнаружения неизвестных атак.

Последнее обеспечивается дополнением множества детекторов, обученных на данных об атаках, некоторым количеством детекторов, сгенерированных случайным образом. Последние не соответствуют ни обучающим данным

нормальной активности, ни обучающим данным об атаках. Если какая-либо строка анализируемых данных соответствует такому детектору, значит, эта строка представляет собой некую аномалию, которая может быть вызвана неизвестными атаками.

Лимфоцит отмечался классом атаки, которой он соответствует. Затем при анализе и выявлении атаки определенным лимфоцитом, класс атаки отмечался в соответствии с классом, обнаруживаемым лимфоцитом. Лимфоциты со случайно сгенерированным детектором, предназначенные для выявления неизвестных атак, в случае выявления отмечали класс атаки как «unknown».

После завершения обучения системой проводился анализ второй, незнакомой половины датасета. С первой же эпохи (под эпохой понимается подача на вход системы всех анализируемых данных один раз) система выявила большинство атак, и сразу же дообучилась на их основе. На следующей эпохе анализа было обнаружено больше атак. И с каждой эпохой уровень ошибок второго рода снижался.

Анализ проводился по различному количеству параметров. Полученный уровень ошибок второго рода в зависимости от количества использованных параметров датасета представлен в Таблице 2.2. Сначала было использовано 16 параметров, результаты анализа оказались приемлемыми, поэтому их число уменьшалось. Как видно из Таблицы 2.2, при использовании 12 параметров и менее наблюдается предел, выделенный серым цветом, при достижении которого от эпохи к эпохе уровень ошибок не снижается, система не может обнаружить некоторые образцы атак по причине недостаточности данных. Уровень ошибок первого рода оставался стабильно низким вследствие высокой эффективности применения алгоритма негативной селекции. Вычислительные эксперименты показали, что его эффективность зависит от количества параметров, а количество эпох влияет на нее незначительно.

Таблица 2.2 – Уровень ошибок второго рода (FNR)

№ эпохи	Количество параметров							
	8	9	10	11	12	13	14	16
1	0,056	0,049	0,047	0,043	0,039	0,029	0,025	0,024
2	0,038	0,035	0,026	0,028	0,019	0,012	0,012	0,011
3	0,038	0,035	0,024	0,024	0,018	0,008	0,008	0,007
4	0,038	0,035	0,024	0,024	0,017	0,007	0,007	0,007
5	0,038	0,035	0,024	0,024	0,017	0,007	0,007	0,006
6	0,038	0,035	0,024	0,024	0,017	0,006	0,006	0,006
7	0,038	0,035	0,024	0,024	0,017	0,006	0,005	0,005
8	0,038	0,035	0,024	0,024	0,017	0,005	0,005	0,004
9	0,038	0,035	0,024	0,024	0,017	0,004	0,004	0,003
10	0,038	0,035	0,024	0,024	0,017	0,003	0,003	0,002

Для наглядности данные Таблицы 2.2 представлены графически на Рисунке 2.10.

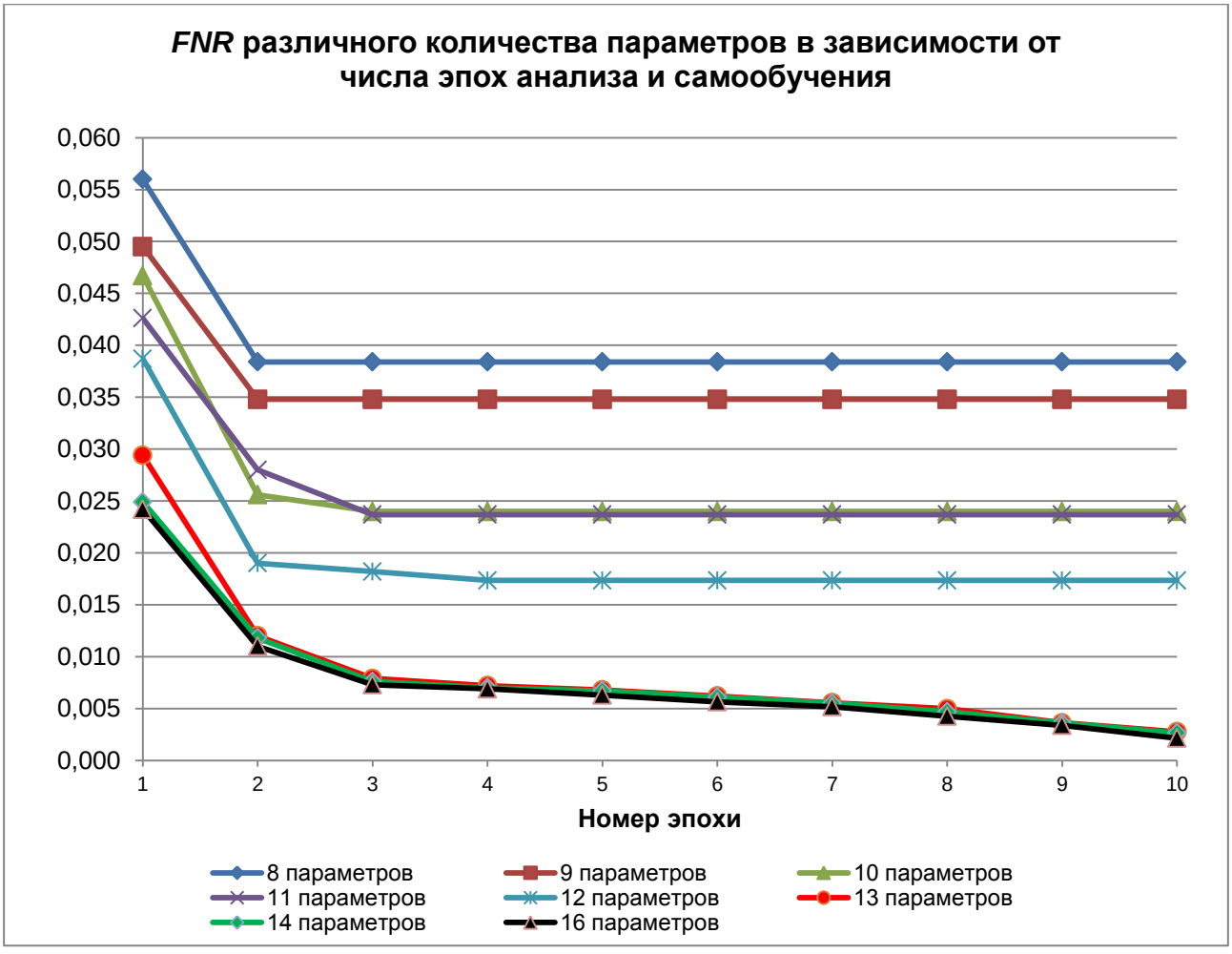


Рисунок 2.10 – Уровень ошибок второго рода (FNR)

Средние показатели уровня ошибок 1-го рода приведены в Таблице 2.3. Таким образом, рационально использовать 13 параметров, если здесь и существует аналогичный предел снижения ошибок, то его уровень менее 0,3%, что является приемлемым, увеличение количества параметров лишь увеличивает затраты вычислительных ресурсов.

Таблица 2.3 – Уровень ошибок первого рода (FPR)

Оцениваемая характеристика	Количество параметров							
	8	9	10	11	12	13	14	16
процент ошибок 1-го рода	0,0012	0,0011	0,0011	0,0010	0,0010	0,0005	0,0005	0,0001

Оценивать эффективность системы корректнее по первой эпохе анализа, полученные показатели представлены в Таблице 2.4.

Таблица 2.4 – Показатели эффективности ИИС после первой эпохи в первой серии экспериментов

Классификатор	FNR	FPR	TNR	TPR (Recall)	Precision	Accuracy	F_1 score
ИИС	0,029	<0,001	0,999	0,970	0,999	0,986	0,985

Уровень ошибок второго рода составил 0,029, что является довольно низким, но недостаточно. Однако большим преимуществом системы, перекрывающим данный недостаток, является то, что после первой же эпохи анализа по 13 параметрам система самообучается и FNR снижается более чем в два раза и составляет уже 1,2%. Следует отметить, что датасет не был предварительно сбалансирован, из него не удалялись атаки, представленные в малом количестве, но, тем не менее, они выявлялись ИИС. Результаты аналогичного эксперимента, проведенного на сбалансированном NSL-KDD, опубликованы в [189], где получены меньшие значения ошибок, но также подтверждена целесообразность использования 13 параметров. Однако

остановимся на результатах текущего эксперимента на несбалансированном NSL-KDD. В итоге, с течением времени система самообучается и выявляет всё больше атак, через некоторое время (10 эпох анализа) ИИС демонстрировала уровень эффективности, показатели которого представлены Таблицей 2.5.

Таблица 2.5 – Показатели эффективности ИИС, достигнутые в результате самообучения в процессе анализа в первой серии экспериментов

Классификатор	FNR	FPR	TNR	TPR (Recall)	Precision	Accuracy	F_1 score
ИИС	0,003	<0,001	0,999	0,997	0,999	0,998	0,998

Также была проведена вторая серия экспериментов по оценке эффективности подсистемы лимфоцитов при обнаружении неизвестных для нее атак. ИИС первично обучалась на 50% данных о нормальной активности, не обучалась на основе данных об атаках, то есть каждая атака, содержащаяся в NSL-KDD, была для нее незнакомой.

Далее исследовалось самообучение системы уже непосредственно в процессе проводимого ею анализа. На вход системы подавалась вторая половина данных о нормальной активности для выявления ошибок первого рода, а также весь набор данных об атаках. Анализ проводился по 13 выбранным параметрам. Ошибки первого рода были аналогично на низком уровне, независимо от числа эпох. Количество ошибок второго рода в зависимости от количества эпох анализа и самообучения представлено в Таблице 2.6.

Конечно, с первой же эпохи система оказалась неспособна выявить сразу большинство атак, поскольку детекторами не перекрывается всё пространство параметров аномалий, так как это потребовало бы значительных ресурсов. Но уже с первой эпохи система обнаружила несколько атак, в том числе те, которые представлены в NSL-KDD в малом количестве, самообучилась на их основе и на следующей эпохе анализа она обнаружила уже больше атак.

Таблица 2.6 – Результаты обнаружения неизвестных атак

Эпоха анализа	Уровень обнаружения атак (Recall)	Ошибки 2-го рода (FNR)
1	0,002	0,998
3	0,016	0,984
5	0,720	0,280
7	0,865	0,135
9	0,964	0,036
11	0,979	0,021
13	0,984	0,016
15	0,990	0,010
17	0,995	0,005
20	0,995	0,005

Значения показателей эффективности ИИС, достигнутые после 20 эпох анализа при её обучении только на данных о нормальной сетевой активности, представлены в Таблице 2.7.

Таблица 2.7 – Значения показателей эффективности ИИС, достигнутые во второй серии вычислительных экспериментов

Классификатор	FNR	FPR	TNR	TPR (Recall)	Precision	Accuracy	F_1 score
ИИС	0,005	<0,001	0,999	0,995	0,999	0,997	0,997

В Таблице 2.8 представлены дополнительные сведения о проведенных экспериментах.

Таблица 2.8 – Дополнительные сведения о проведенных экспериментах

Характеристика	Эксперимент 1	Эксперимент 2
Особенность выявления атак	ИИС: – классифицирует известные атаки, нормальное взаимодействие, – выявляет неизвестные атаки	все выявляемые атаки неизвестны ИИС
Процессор	AMD Athlon(tm) X4 870K Quad Core 3,90GHz	
Расходуемый объём ОЗУ	136 МБ	
Объем обучающей выборки	33671 образцов нормальной сетевой активности; 29315 образцов атак	33671 образцов нормальной сетевой активности;

Продолжение таблицы 2.8

Характеристика	Эксперимент 1	Эксперимент 2
Количество детекторов (лимфоцитов)	500 000	
Количество создаваемых клонов	3	
Количество изменяемых параметров при создании клонов	2	
Аффинность	12	
Время, затраченное на первичное обучение системы	6 часов	8 часов
Время, затраченное на достижение результатов, представленных в таблицах 2.5 и 2.7, посредством самообучения	25 часов	40 часов

Таким образом, подсистема лимфоцитов успешно обнаруживает как известные, так и неизвестные для нее атаки, и эффективно обучается на их основе. Первичные значения показателей эффективности при первой «встрече» ее с конкретным экземпляром атаки могут быть не самыми высокими, однако способность системы самообучаться «на лету» позволяет ей резко увеличивать свою эффективность, адаптируясь к обнаружению актуальных для объекта атак.

После первичного анализа подсистемой лимфоцитов, данные об атаках передавались подсистеме ДК. В данном эксперименте ДК анализировали только поступающие от лимфоцитов данные без подключения дополнительных источников, таких как SIEM и пр. Если обнаруживалась известная атака (отличная от класса «unknown»), система устанавливала уровень опасности выше порогового допустимого уровня. То есть ни одна известная атака не расценивалась как безопасная и не игнорировалась.

В этом случае системой оценивался общий уровень опасности: чем чаще или продолжительнее обнаруживаются данные, характерные для известных атак, или, более того, одной и той же атаке, тем выше становился уровень опасности. В рассматриваемой ситуации подсистема ДК никак не влияла на точность распознавания, но оценивала уровень опасности.

Обнаружение атак класса «unknown» обычно происходило, когда система лимфоцитов была обучена только на данных о нормальном сетевом

взаимодействии и все атаки расценивались как неизвестные. В таком случае уровень опасности от каждой атаки поднимался незначительно, и если анализ был растянут во времени, то не все атаки сразу вызывали превышение порогового уровня опасности, после которого выполняется оповещение об опасности. Пороговый уровень опасности выбирается экспертным путем или на основе результатов вычислительных экспериментов для конкретного объекта защиты.

Некоторые атаки, особенно представленные в малом количестве или обнаруженные первыми, лишь незначительно повышали уровень опасности и оставались бы без внимания, но обнаружение последующих атак обеспечивало превышение порогового уровня опасности.

Рассматривать результаты общего анализа не имеет особого смысла, так как точность обнаружения той или иной атаки зависит от очередности ее возникновения, поэтому для наглядности настройки чувствительности порога опасности каждый класс атак был проанализирован отдельно. Результаты анализа представлены в Таблице 2.9, где t – значение порога опасности.

Таблица 2.9 – Результат работы алгоритма Дендритных клеток

Вид атаки	Количество примеров атаки, поданных на вход системы	Достигнутый уровень опасности	Доля атак, не вызвавших реакцию			
			t = 5	t = 50	t = 500	t = 5000
1	*2*	*3*	*4*	*5*	*6*	*7*
neptune	41020	1682640400	0,0%	0,0%	0,0%	0,1%
satan	3616	13075456	0,1%	0,2%	0,1%	0,2%
ipsweep	3582	12830724	0,1%	0,2%	0,6%	2,0%
portsweep	2917	8508889	0,1%	0,2%	0,6%	2,0%
smurf	2634	6937956	0,1%	0,3%	0,8%	2,4%
nmap	1486	2208196	0,1%	0,5%	0,8%	2,7%
back	952	906304	0,2%	0,7%	1,5%	4,8%
teardrop	888	788544	0,2%	0,8%	2,3%	7,5%
warezclient	886	784996	0,2%	0,8%	2,5%	8,0%
pod	200	40000	1,0%	3,5%	2,5%	8,0%
guess_passwd	53	2809	3,8%	13,2%	11,0%	35,5%
buffer_overflow	30	900	6,7%	23,3%	41,5%	100,0%
warezmaster	20	400	10,0%	35,0%	73,3%	100,0%
land	18	324	11,1%	38,9%	100,0%	100,0%
imap	11	121	18,2%	63,6%	100,0%	100,0%

Продолжение таблицы 2.9

1	*2*	*3*	*4*	*5*	*6*	*7*
rootkit	10	100	20,0%	70,0%	100,0%	100,0%
loadmodule	9	81	22,2%	77,8%	100,0%	100,0%
ftp_write	8	64	25,0%	87,5%	100,0%	100,0%
multihop	7	49	28,6%	100,0%	100,0%	100,0%
phf	4	16	50,0%	100,0%	100,0%	100,0%
perl	3	9	66,7%	100,0%	100,0%	100,0%

Увеличение порога опасности увеличивает долю игнорируемых неизвестных атак. Это может расцениваться как недостаток системы, однако не все аномалии являются атаками. Датасеты, как правило, содержат образцы атак и образцы нормального сетевого трафика, но не содержат образцы неопасных аномалий в достаточном количестве. Они, возможно, присутствуют в образцах нормального трафика, но далеко не все.

Если использовать только подсистему лимфоцитов ИИС, которая продемонстрировала высокую точность и адаптацию, она обнаружит и неопасную аномалию, обучится на ее основе и, возможно, обнаружит подобные. Да, в классификации «свой/чужой» это не будет ошибкой 2-го рода, аномалия действительно чужая, но острое реагирование на любую аномалию характерно «параноидальному» уровню защиты, что приводит к большому числу оповещений и к наиболее вероятному их игнорированию со стороны администратора безопасности. Применение предложенного алгоритма ДК позволяет экспертным путем снизить чувствительность системы, определить необходимую частоту и количество возникающих разнородных или, наоборот, подобных аномалий для их классификации в качестве атаки и оповещения администратора.

Тогда возникает следующая проблема. ДК не будут препятствовать обнаружению неизвестных атак, для которых характерно повторяющееся воздействие, но если атака характеризуется только единожды взятой аномалией, она может быть проигнорирована ДК. Но, во-первых, при параноидальном мониторинге аномалий она в любом случае, скорее всего, просто затеряется среди

сотен других оповещений, а во-вторых, такие атаки должны выявляться за счет анализа дендритными клетками других подключенных источников.

Нужно выполнить настройку так, чтобы если внешние подключаемые системы, такие как SIEM, сообщают о нехарактерных состояниях или действиях, поднимающих общий уровень опасности, и возникает аномалия, то уровень опасности превысит пороговое значение. Но настройка правил и чувствительности должна выполняться отдельно для каждого объекта.

Таким образом, рассмотренные алгоритмы взаимно интегрируемы, служат для обнаружения атак и опасностей в составе двухуровневой ИИС. Подсистема лимфоцитов осуществляет параноидальное обнаружение всех атак, в том числе неизвестных, самостоятельно обучается на их основе лучшему выявлению атак, подобных обнаруженным. Подсистема ДК повышает общий уровень опасности и сигнализирует об этом при обнаружении известной атаки, а также при обнаружении большого числа аномалий или при получении информации об опасных состояниях из внешних систем, а также при совокупности событий с настраиваемой чувствительностью и набором правил.

В целом, предложенная ИИС представляет собой двухуровневую систему принятия решений, где на нижнем уровне принимается решение о наличии атаки или неизвестной аномалии, на верхнем производится принятие решение о текущем уровне опасности с учетом анализа данных от подключаемых систем и о необходимости применения контрмер. Двухуровневая ИИС построена в классе распределенных многоагентных систем, где агенты нижнего уровня распределяются по подсетям и анализируют сетевой трафик на предмет выявления известных и неизвестных атак и аномалий, а агенты верхнего уровня выполняют граничные вычисления, собирая и анализируя данные, полученные от всех низкоуровневых агентов соответствующей сети, а также подключаемых других систем.

2.4 Многоагентная распределенная система обнаружения атак и аномалий сетевого трафика

Разработана распределенная многоагентная двухуровневая архитектура ИИС, схема реализации которой представлена на Рисунке 2.11, где ИД – источники данных.

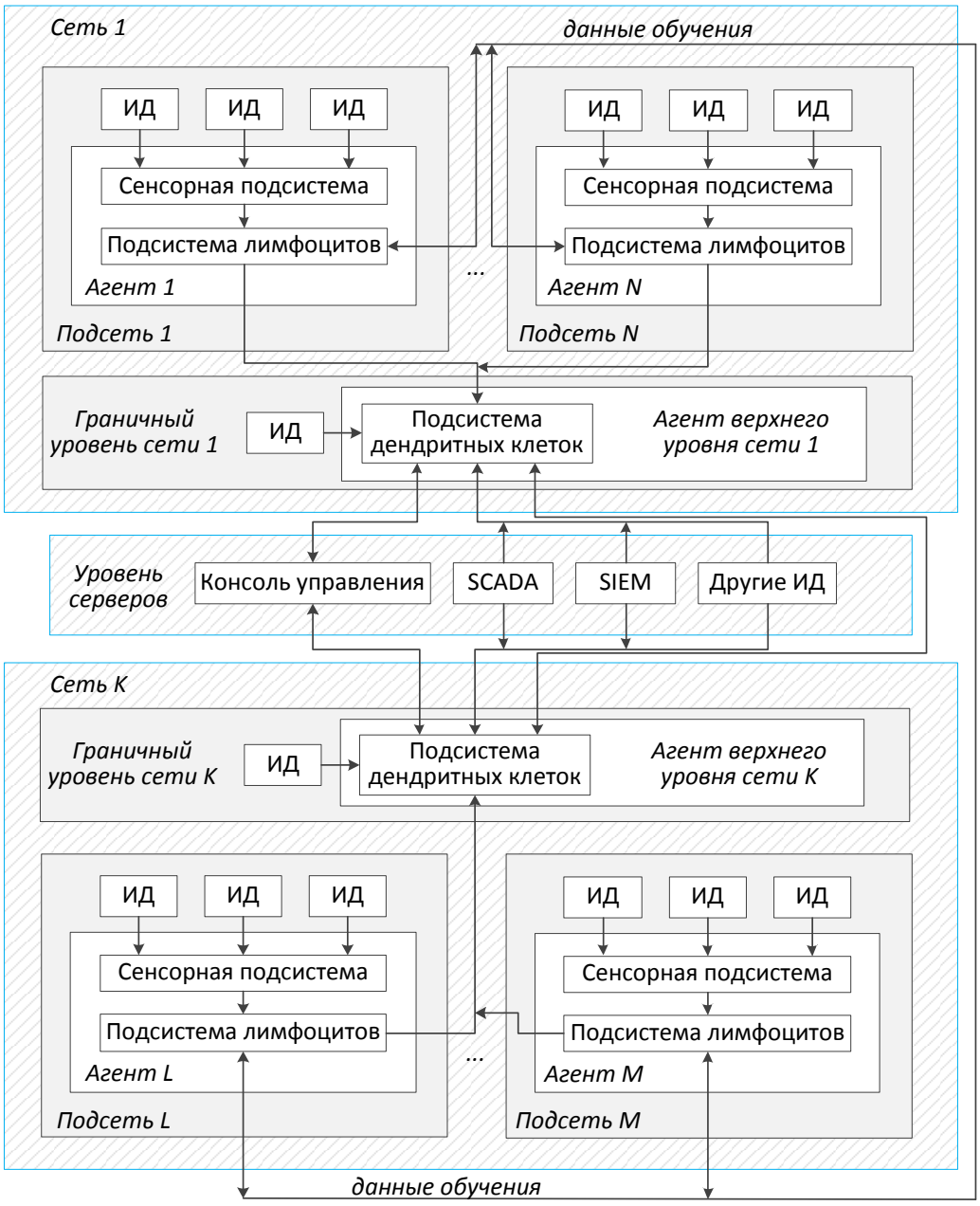


Рисунок 2.11 – Схема реализации многоагентной распределенной двухуровневой ИИС

Агенты нижнего уровня распределены по подсетям, выполняют анализ посредством подсистем лимфоцитов, взаимодействуют друг с другом, взаимно обучают друг друга, выявляют атаки и аномалии. На верхнем уровне системы функционируют агенты, реализующие алгоритм ДК, они располагаются на границе сети, выполняют граничные вычисления, взаимодействуют друг с другом.

Агенты нижнего уровня, в случае выявления неизвестной аномалии, направляют всем другим агентам нижнего уровня во всех сетях информацию о выявившем ее лимфоците. Другие агенты осуществляют клональную селекцию данного лимфоцита и тем самым обучаются лучшему выявлению подобных аномалий, даже несмотря на то, что никогда с ней не встречались.

Следует отметить, что перед эксплуатацией системы, даже уже обученной на датасетах, следует собрать образцы сетевого трафика, включая трафик взаимодействия между агентами, чтобы он не расценивался системой как аномальный.

Существует вероятность того, что злоумышленник сможет имитировать сообщения, содержащие информацию о клонируемом лимфоците, и попробует отправить лимфоцит, реагирующий на нормальную сетевую активность, агентам системы для формирования ошибок первого рода, но все входящие лимфоциты и их клоны также подвергаются негативной селекции, и такой лимфоцит будет отброшен системой.

С другой стороны, злоумышленник может реализовать DoS-атаку на агент нижнего уровня, отправляя ему поток лимфоцитов, и агент вынужден будет обрабатывать каждый из них, проводя негативную и клональную селекцию.

В этом случае разумна реализация электронной подписи данных или их шифрование с контрольной суммой на pre-shared ключах. Криптографические механизмы в данной работе не рассматриваются.

В рамках данной работы эта проблема частично решена реализацией негативной и клональной селекции в потоке, отдельном от процедуры анализа. Даже если предстоит обработка большого объема входящих лимфоцитов, поток

зависает, не успевает обрабатывать входящие фиктивные данные, не может вовремя обрабатывать корректные реальные обучающие данные от других агентов, но всё же анализ в отдельном потоке не прекратится и система продолжит обнаруживать атаки и аномалии. Да и необычно большой поток данных о среагировавших лимфоцитах сам по себе является аномалией и, вероятнее всего, будет обнаружен и значительно увеличит уровень опасности, определяемый дендритными клетками.

Агенты верхнего уровня, реализующие алгоритм ДК, устанавливаются на границах каждой сети, выполняют граничные вычисления, анализируют информацию от подсистем лимфоцитов, граничных маршрутизаторов, МЭ, получают информацию от различных источников из Центра обработки данных (ЦОД), анализируют уровень опасности, сигнализируют о нем в консоли управления.

Выводы по второй главе

1. Для решения задачи мониторинга ИБ сетей промышленного Интернета вещей предложено применение многоуровневого интеллектуального анализа данных сетевого трафика, где на нижних двух уровнях работает распределенная двухуровневая ИИС, на верхнем – система классификации событий ИБ.

2. Рассмотрен алгоритм выбора и нормализации данных сетевого трафика на основе анализа параметров сетевых соединений, используемых в датасетах, предназначенных для построения и тестирования систем обнаружения атак, основанный на приведении исходных значений параметров к заданному целочисленному диапазону, разделении датасета на подмножества данных об атаках и данных о нормальном сетевом взаимодействии, определении частоты совпадений параметров между подмножествами, ранжировании по наименьшему проценту совпадений (наибольшей информативности), проведении

вычислительных экспериментов по обнаружению атак с использованием различного количества наиболее информативных параметров.

3. Подсистема лимфоцитарного анализа осуществляет обнаружение всех атак, в том числе неизвестных, самостоятельно обучается на их основе лучшему выявлению атак, подобных обнаруженным. Подсистема ДК повышает общий уровень опасности и сигнализирует об этом при обнаружении известной атаки, а также при обнаружении большого числа аномалий или при получении информации об опасных состояниях из внешних систем, а также при совокупности событий с настраиваемой чувствительностью и набором правил.

4. Представлены результаты реализации двухуровневой ИИС в классе распределенных многоагентных систем, где агенты нижнего уровня распределяются по подсетям и анализируют сетевой трафик на предмет наличия известных или неизвестных атак и аномалий, а агенты верхнего уровня выполняют граничные вычисления, собирая и анализируя данные со всех низкоуровневых агентов соответствующей сети, а также подключаемых других систем. Проведены вычислительные эксперименты с использованием нормализованного набора данных NSL-KDD. Вычислительные эксперименты показали высокую эффективность применения двухуровневой ИИС.

3 Разработка и исследование алгоритмов мониторинга информационной безопасности промышленного Интернета вещей с использованием гибридных технологий искусственных иммунных систем и методов машинного обучения

3.1 Концепция построения гибридной распределенной интеллектуальной системы мониторинга сетевых атак на системы промышленного Интернета вещей

Концепция построения гибридной распределенной интеллектуальной системы мониторинга (РИСМ) атак и аномалий сетевого трафика, структура которой представлена на Рисунке 3.1, предполагает обеспечение «глубокой защиты» (Defense-in-Depth) промышленного Интернета вещей.

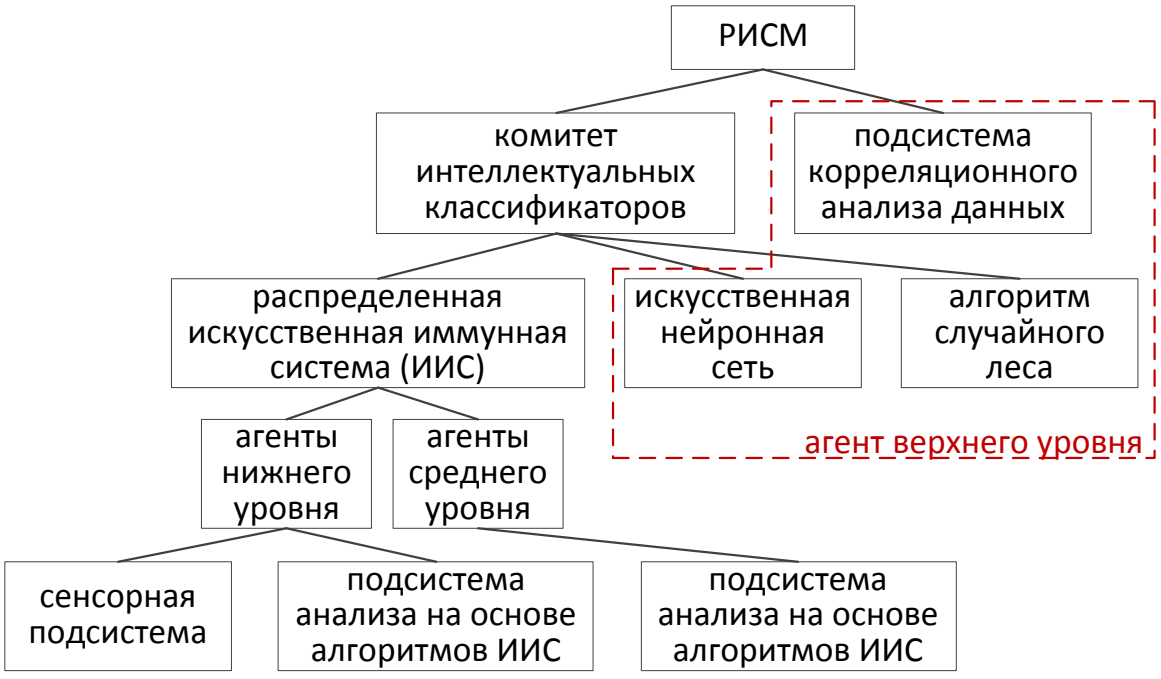


Рисунок 3.1 – Структура РИСМ

Положения предлагаемой концепции построения гибридной РИСМ атак и аномалий сетевого трафика (далее – концепции построения РИСМ):

1. В качестве входных данных эта система предполагает использование разнородной информации с большого числа распределенных датчиков (сенсоров), собирающих данные о сетевом трафике IIoT, представленных на Рисунке 3.1 как сенсорная подсистема, и внешних систем, регистрирующих возможное появление событий и инцидентов ИБ в различных точках на различных уровнях IIoT.

2. Предлагаемая система имеет гибридную архитектуру, подробно представленную в разделе 4.1. (на Рисунке 4.1), реализующую комплекс методов (технологий) искусственного интеллекта, статистической обработки (корреляционного анализа) данных, построена с применением многоагентного подхода.

3. Центральное место в составе РИСМ отводится представленной в виде множества взаимодействующих агентов интеллектуальной системе обнаружения атак и аномалий, построенной на основе двухуровневой искусственной иммунной системы (ИИС), функционирующей, в свою очередь, на основе комплекса алгоритмов отрицательного отбора, клональной селекции, алгоритма дендритных клеток и др., реализуемых лимфоцитами на агентах нижнего уровня, интегрируемых в подсети конечных IIoT-устройств, и алгоритмов, связанных с теорией опасности и дендритными клетками, реализуемых на агентах верхнего уровня, выполняющих граничные вычисления, определение состояния опасности на основе данных, получаемых с нижнего уровня и внешних систем, таких как межсетевые экраны, сетевое оборудование, SIEM, SCADA и др.

4. Интегрирование в серверную компоненту РИСМ на верхнем уровне системы классификации событий ИБ, реализованной в виде комитета классификаторов с использованием таких методов машинного обучения, как алгоритм случайного леса (СЛ) и искусственная нейронная сеть (ИНС), позволяет обеспечить более высокую точность обнаружения и распознавания атак и аномалий сетевого трафика IIoT. Многоагентная архитектура проектируемой системы обеспечивает её масштабируемость и простоту интеграции в системы IIoT.

5. Агенты двухуровневой ИИС осуществляют анализ трафика, выявляют аномалии, самообучаются на их основе, определяют уровень опасности, взаимодействуют с другими агентами ИИС в рамках взаимного обучения и поддержания активности детекторов в соответствии с принципами идиотипической сети. Они передают данные об угрозах в ЦОД, где серверная компонента осуществляет дополнительный анализ на основе совокупности мнений (решений) ИИС, ИНС и СЛ, принимает окончательное решение о наличии той или иной атаки или аномалии, степени ее критичности, оповещает об инцидентах.

6. Данная система позволяет в будущем интегрировать функционал автоматизации ТП. К примеру, для беспроводных сенсорных сетей при обнаружении одного экземпляра атаки на TDMA-расписание протокола LEACH, подсистеме реагирования передается информация о необходимости отправки на соответствующую подсеть сигнала, инициирующего новое формирование расписания. Если из данной подсети поступает информация о повторном подобном инциденте, то подаётся сигнал об отключении потенциально скомпрометированного узла от главы кластера; система правил редактируется в консоли администратора.

7. Интегрированная в РИСМ подсистема корреляционного анализа данных (КАД) позволяет выявлять неочевидные значимые взаимосвязи между событиями ИБ. Её графический интерфейс позволяет удобно оценить общий уровень защищенности, выделить наиболее значимые события, определить приоритеты в реагировании. Эта подсистема также помогает принять наиболее рациональное решение для автоматического реагирования. Вычислительные эксперименты показали высокую эффективность предложенной системы [191].

Таким образом, предложенная концепция построения гибридной распределенной интеллектуальной системы мониторинга (РИСМ) атак и аномалий сетевого трафика промышленного Интернета вещей базируется на интеграции нескольких взаимодействующих между собой систем – распределенной многоагентной двухуровневой ИИС, системы классификации

событий ИБ в виде комитета классификаторов и подсистемы корреляционного анализа данных (КАД) об инцидентах ИБ, в том числе полученных от внешней SIEM-системы.

3.2 Механизмы взаимодействия гибридной интеллектуальной системы обнаружения атак и сетевых аномалий с подсистемой корреляционного анализа событий ИБ (менеджмента инцидентов ИБ) SIEM-системы

Если сеть предприятия, на котором реализуется РИСМ обнаружения атак и аномалий, подключена к существующей SIEM-системе, то осуществляется непосредственная передача этой системе данных о выявленных атаках на сети промышленного Интернета вещей. А также, при соответствующей настройке SIEM-системы, получение результатов корреляционного анализа в РИСМ от SIEM-системы. При этом необходимо дополнительно настроить существующую SIEM-систему для передачи оповещений об инцидентах ИБ агентам подсистемы дендритных клеток РИСМ в соответствии с анализируемыми ими сегментами сети.

Одной из наиболее известных SIEM-систем с открытым исходных кодом является OSSIM (Open Source Security Information Management) [165]. Штатный комплект данной системы включает [164]:

– подсистему сбора, анализа и корреляция событий – SIEM;

– хостовую систему обнаружения вторжений (HIDS) – OSSEC;

– сетевую систему обнаружения вторжений (NIDS) – Suricata;

– беспроводную систему обнаружения вторжений (WIDS) – Kismet;

– мониторинг узлов сети – Nagios;

– подсистемы анализа сетевых аномалий – P0f, PADS, FProbe, Arpwatch и др.;

– сканер уязвимостей – OpenVAS;

– систему обмена информацией об угрозах между пользователями OSSIM – OTX;

– более 200 плагинов для парсинга и корреляции логов с множества внешних устройств и служб.

Наиболее простой способ интеграции OSSIM с предлагаемой системой – настройка отправки оповещений об инцидентах по почтовым протоколам от OSSIM к РИСМ. Описание данной настройки со стороны OSSIM приведено в [164].

В качестве примера отработки взаимодействия РИСМ с SIEM-системой в рамках вычислительных экспериментов, а также для функционирования в условиях отсутствия сторонней SIEM-системы, была разработана встроенная в РИСМ подсистема корреляционного анализа данных (КАД) о событиях ИБ, которая получает данные об обнаруженных атаках после их обработки комитетом классификаторов, представленном на Рисунке 3.1 и в пункте 4 концепции построения РИСМ, изложенной в разделе 3.1, и сравнивает их со всеми предыдущими атаками данного класса.

В основу построения подсистемы КАД положен метод, реализующий вычисление коэффициентов корреляции Пирсона между событиями ИБ. Подобный подход описан в [146]:

1. Вычисление частных коэффициентов корреляции Пирсона по группам событий определенного временного окна. Осуществляется на основе двух показателей: временной задержки между двумя сравниваемыми событиями в секундах – dT и относительного веса w_i прямых связей между анализируемыми событиями, определяемого по формуле:

$$w_i = \frac{N_i^{iden}}{N_i^{direct}}, \tag{3.1}$$

где N_i^{iden} – число идентичных элементов 2-х анализируемых векторов, N_i^{direct} – размерность векторов-признаков события.

Частные коэффициенты корреляции между двумя событиями рассчитываются по формуле:

$$r_{priv} = \frac{n\sum_{i=1}^{n-1} w_i dT_i - \sum_{i=1}^{n-1} w_i \times \sum_{i=1}^{n-1} dT_i}{\sqrt{n\sum_{i=1}^{n-1}(w_i)^2 - \left(\sum_{i=1}^{n-1} w_i\right)^2} \times \sqrt{n\sum_{i=1}^{n-1}(dT_i)^2 - \left(\sum_{i=1}^{n-1} dT_i\right)^2}}, \quad (3.2)$$

где n – число произошедших событий, w_i – относительный вес прямых связей между двумя событиями, dT_i – разница во времени между двумя событиями.

Отметим, что предыдущая формула, согласно данным [61], является результатом применения формул выборочных средних и раскрытия сумм для исходной формулы расчета коэффициента Пирсона:

$$r_{priv} = \frac{\sum_{i=1}^{n-1}(w_i - \bar{w})(dT_i - \overline{dT})}{\sqrt{\sum_{i=1}^{n-1}(w_i - \bar{w})^2} \times \sqrt{\sum_{i=1}^{n-1}(dT_i - \overline{dT})^2}}. \quad (3.3)$$

2. Общий коэффициент корреляции между типами событий рассчитывается как среднее арифметическое частных коэффициентов корреляции [146].

Анализ показал, что подход, предлагаемый в [146], обеспечивает высокие значения коэффициента корреляции, если с увеличением количества совпадающих свойств событий увеличивается и временной интервал между событиями. То есть коэффициент корреляции по группе событий будет высоким, если произошло событие *X* и соблюдаются условия:

– большое количество свойств события *X* совпадает с соответствующими свойствами Y_0 – первого события из анализируемого множества; между событиями *X* и Y_0 соответственно прошло много времени;

– чуть меньшее количество свойств события *X* совпадает с соответствующими свойствами Y_1 – второго события из анализируемого множества; соответственно между событиями *X* и Y_1 прошло чуть меньше

времени, и так далее; для каждого последующего события *Y* уменьшается не только временной интервал с событием *X*, но и количество совпадающих свойств.

Это приводит к тому, что внутри группы новое событие ИБ, полностью соответствующее по своим свойствам самому давнему событию в пределах временного окна, будет иметь высокий коэффициент корреляции, даже если по своим совпадающим свойствам оно сильно отличается от самого недавнего события.

Но в прямой зависимости между числом совпадающих свойств и временным интервалом кроется и существенный недостаток. Если возникнет событие ИБ, полностью соответствующее последнему событию группы, то количество совпадающих свойств будет большим, временной интервал маленьким, при этом коэффициент корреляции резко уменьшится, хотя связь этих событий очевидна.

В рамках РИСМ реализована корреляция на основе обратной пропорциональности. Коэффициент корреляции повышается, если при росте числа совпадений свойств уменьшается временной интервал между событиями. Здесь возникает обратный недостаток: если новое событие ИБ по свойствам полностью совпадает с первым событием множества, временной интервал между ними будет большим и коэффициент корреляции уменьшится. Однако гораздо менее критичен случай, когда не считаются коррелирующими события ИБ, находящиеся далеко друг от друга во времени, нежели близко. Также возможно подключение к подсистеме корреляционного анализа дополнительных источников данных и анализ информации на основе правил в формате, представленном в [47].

Работа подсистемы корреляционного анализа данных (КАД) тестировалась ранее в непосредственном взаимодействии с распределенной ИИС, на тот момент без реализации комитета классификаторов согласно схеме, изображенной на Рисунке 3.2.

Полученные результаты первичного тестирования взаимодействия ИИС и подсистемы КАД опубликованы: в [33] – на основе набора данных NSL-KDD и в [191] – на основе датасета WSN-DS.

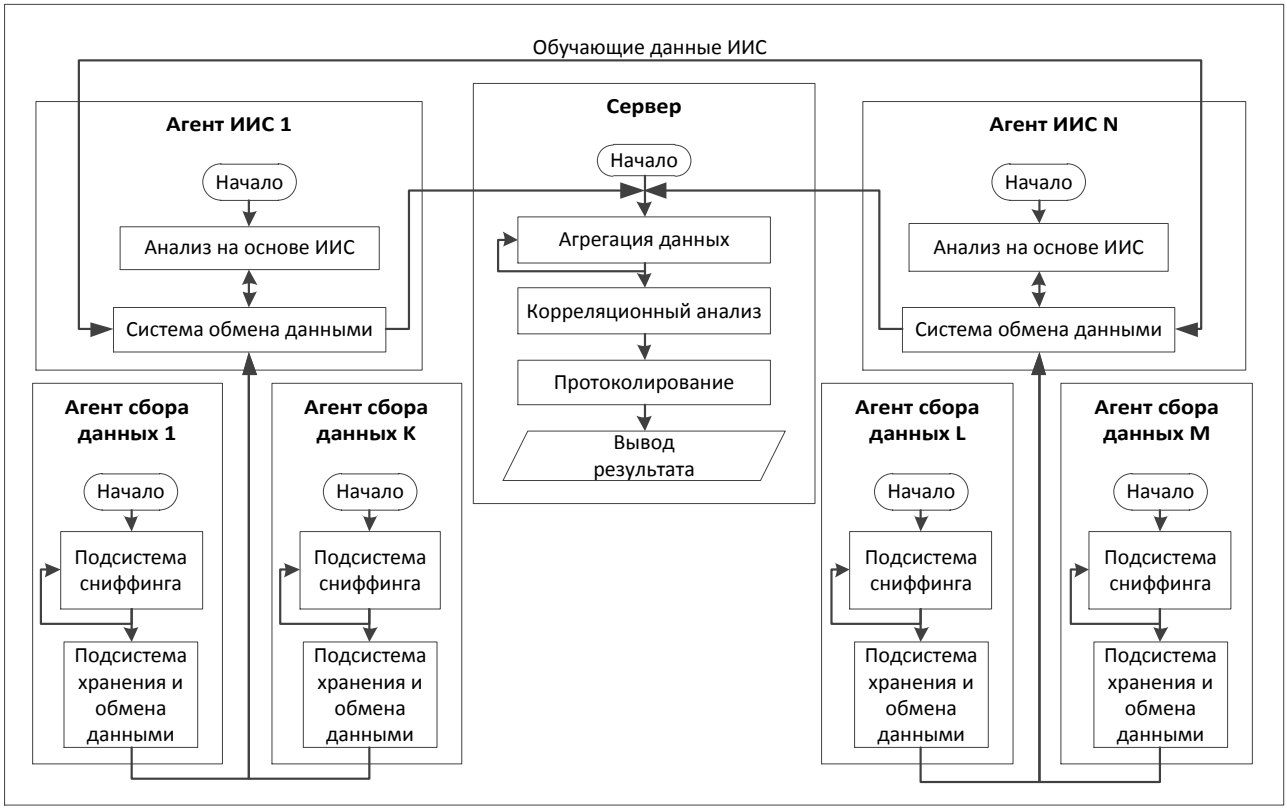


Рисунок 3.2 – Схема взаимодействия ИИС и подсистемы КАД

Пример формируемого отчета на основе WSN-DS представлен на Рисунке 3.3. Система анализирует данные по каждой группе инцидентов ИБ отдельно, а также по всем инцидентам совместно в рамках заданного временного интервала. Подробное описание проведенных вычислительных экспериментов и полученных результатов приведено в [191].

Но в рамках РИСМ подсистема КАД функционирует не только как самостоятельная подсистема, получающая данные от распределённой ИИС, анализирующая их, выводящая результат анализа на консоль администрирования, но подсистема КАД также взаимодействует с подсистемой дендритных клеток, передавая ей текущие данные об инцидентах, анализируемых дендритными клетками как сигналы опасности.

Также результаты анализа конкретной атаки или аномалии подсистемой КАД объединяются с результатами анализа этой атаки или аномалии комитетом классификаторов, выводится уже совместный результат.

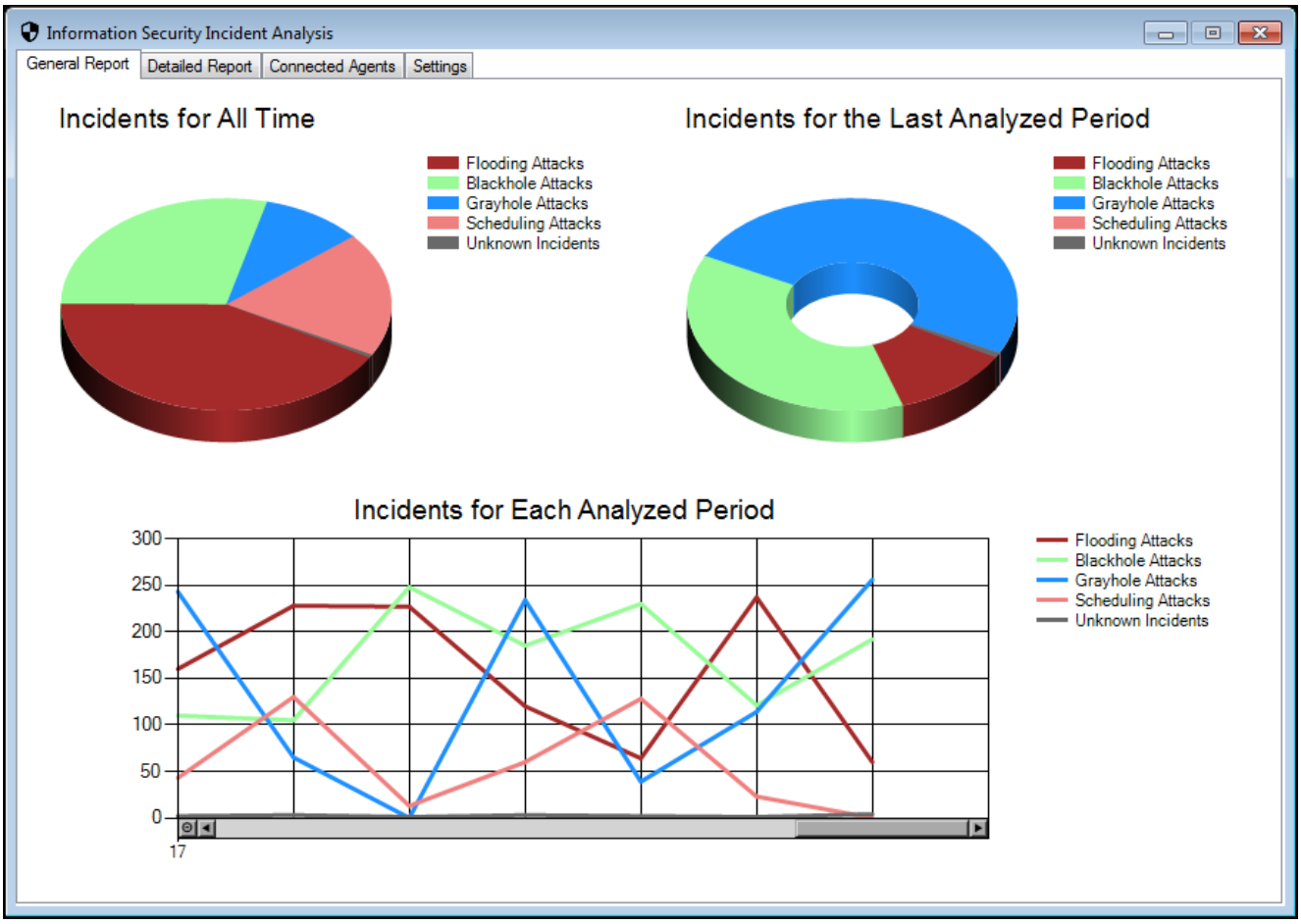


Рисунок 3.3 – Отчет, формируемый подсистемой корреляционного анализа инцидентов ИБ (опубликован в [191])

Таким образом, данные о событии ИБ, например, об атаке и её классе, подаются на вход подсистемы корреляционного анализа, где осуществляется их сопоставление со всеми ранее зарегистрированными атаками данного класса. В случае выявления корреляционной связи между событиями, вычисляется общее количество коррелирующих между собой атак данного класса, прямо пропорционально определяется их критичность.

Результаты передаются подсистеме объединения результатов анализа, а также соответствующим агентам двухуровневой ИИС для корректировки уровня опасности.

Подключение подсистемы КАД описанным образом позволяет принимать итоговое решение не только на основе результатов классификации, полученных комитетом, но и учитывать уровень значимости события ИБ, определяемого рассматриваемой подсистемой, определить приоритеты в реагировании.

Таким образом, интеграция в состав РИСМ обнаружения атак и аномалий подсистемы корреляционного анализа позволяет дополнительно выделить наиболее критичные атаки и инциденты ИБ, определить приоритеты в реагировании, скорректировать уровень опасности, используемый в двухуровневой ИИС на этапе анализа, в том числе с использованием данных от сторонних источников (SIEM-систем).

3.3 Гибридная интеллектуальная система обнаружения атак и сетевых аномалий на основе искусственной иммунной системы и комитета классификаторов

Разработан алгоритм функционирования предлагаемой двухуровневой РИСМ обнаружения атак и аномалий сетевого трафика промышленного Интернета вещей, блок-схема которого представлена на Рисунке 3.4. На первом уровне (уровне распределенной сети) осуществляется сбор данных о сетевом взаимодействии IIoT-устройств посредством сниффинга, зеркалирования, получения данных трафика.

Затем происходит извлечение анализируемых параметров, их нормализация, после чего выполняется анализ данных посредством двухуровневой ИИС, подробно рассмотренный во второй главе.

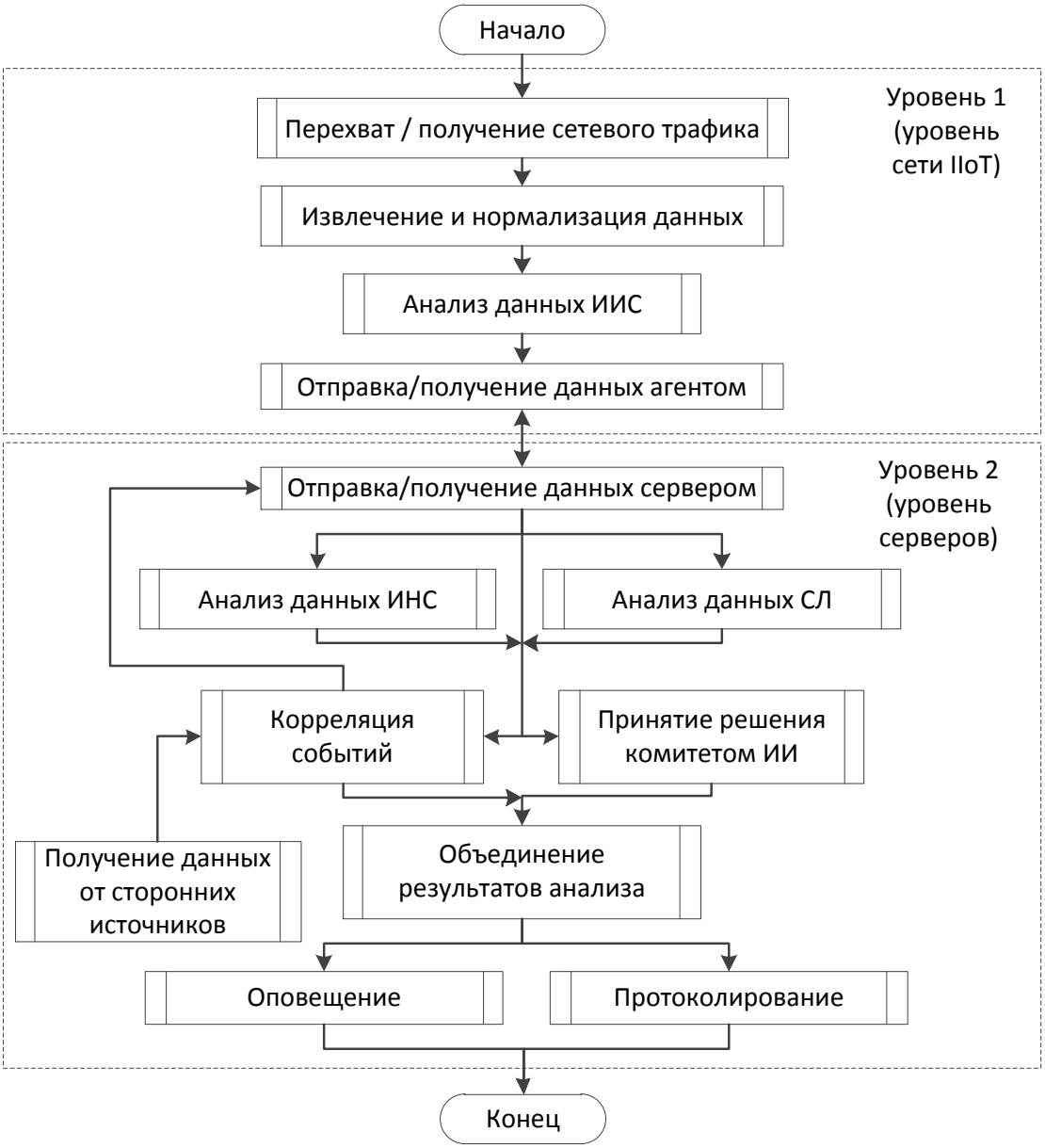


Рисунок 3.4 – Блок-схема алгоритма функционирования предлагаемой двухуровневой РИСМ обнаружения атак и аномалий сетевого трафика IIoT

Данные об атаках, обнаруженных двухуровневой ИИС, отправляются на верхний уровень системы (уровень серверов), где осуществляется агрегация данных, проводится их дополнительный анализ с помощью искусственной нейронной сети (ИНС) и алгоритма случайного леса (СЛ).

На основе выводов трёх классификаторов (ИИС, ИНС, СЛ) принимается решение об отсутствии / или наличии и итоговом классе атаки (аномалии), выполняется корреляционный анализ данных и затем, с учетом выявленного

класса атаки (аномалии) и результатов корреляционного анализа, принимается окончательное решение о критичности атаки (аномалии).

Принятие решения о классе атаки (аномалии) осуществляется с учетом известной теоремы Кондорсе [129] о комитете экспертов (присяжных), согласно которой, если компетентность экспертов выше 0,5, то увеличение числа экспертов приводит к повышению точности результата. Согласно [122], уровень ошибок комбинации алгоритмов всегда гарантированно ниже средней ошибки этих алгоритмов.

Рассмотрим подробнее различные схемы голосования и принятия решений при использовании комитета классификаторов. Аналогичная задача обсуждается в стандарте ГОСТ Р 54411-2018/ISO/IEC TR 24722:2015 «Информационные технологии. Биометрия. Мультимодальные и другие мультибиометрические технологии» [2]. В стандарте выделено объединение схем принятия решений (классификаторов) на следующих уровнях:

– принятие итогового решения основывается на уже принятых локальных решениях;

– принятие итогового решения основывается на результатах сравнения образца данных с эталонным значением.

В нашем случае применим первый вариант. В упомянутом стандарте в качестве простого объединения на уровне принятия решений предлагается использовать логические функции «И», «ИЛИ», а также голосование на основе большинства по мажоритарному принципу.

В качестве модифицированной схемы объединения решений на данном уровне предлагается применение:

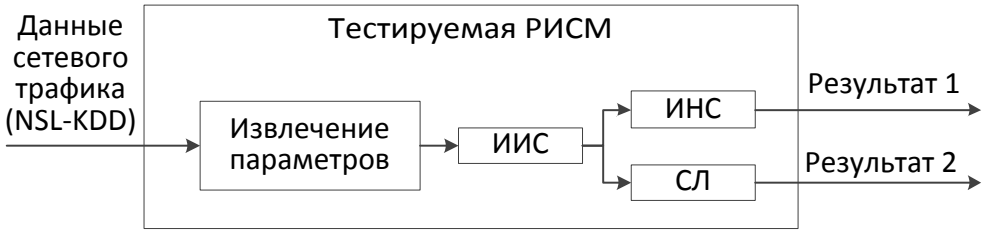
– многоуровневой модели, предполагающей при соответствии результата сравнения критерию первого эксперта, изменение порога для соответствия критерию второго эксперта;

– последовательной модели, предполагающей выполнение сравнения сначала одним экспертом, при достижении определенного критерия и высокой

достоверности сравнение заканчивается, в ином случае, если эта достоверность недостаточная, проводится оценка вторым экспертом и так далее.

Первоначальная гипотеза относительно построения комитета классификаторов заключалась в следующем. Благодаря алгоритму негативной селекции, ИИС предположительно должна демонстрировать один из самых низких уровней ошибок первого рода, к тому же она способна выявлять как известные, так и неизвестные атаки. Но тогда рационально построить систему последовательного принятия решений, где ИИС соответственно выявляет в совокупности известные и неизвестные атаки с низким уровнем ошибок первого рода (False Positives), и если выявлено наличие атаки, то она передается другому классификатору для определения её конкретного класса.

Анализ литературных источников показал, что, как правило, в числе лидеров по эффективности обнаружения атак (без учета ИИС) оказываются такие методы машинного обучения, как искусственная нейронная сеть (ИНС) и алгоритм случайного леса (СЛ). Были проведены вычислительные эксперименты с целью сравнения результатов совместной работы ИИС-ИНС и ИИС-СЛ, как представлено на Рисунке 3.5.



ИИС – искусственная иммунная система;

ИНС – искусственная нейронная сеть; СЛ – случайный лес;

Рисунок 3.5 – Схема взаимодействия классификаторов в первом эксперименте

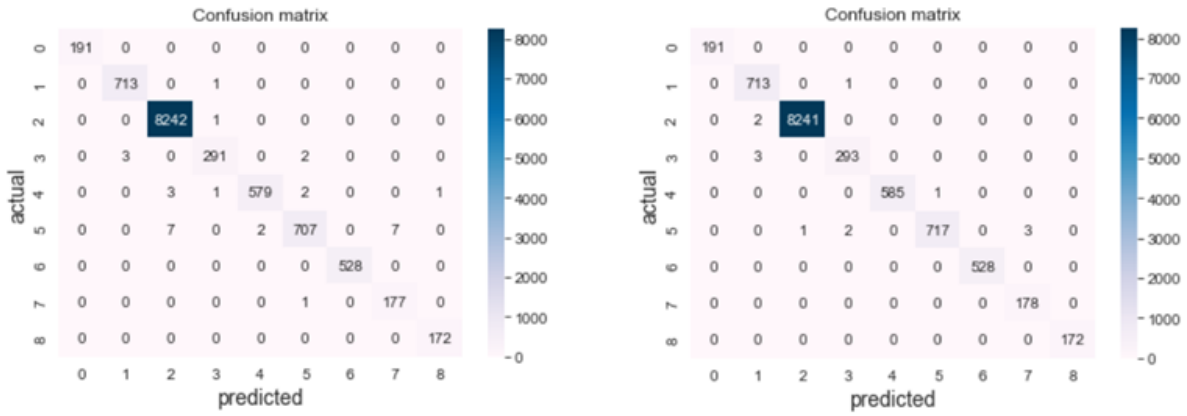
Вычислительные эксперименты проводились на основе датасета NSL-KDD. В отличие от ИИС, ИНС и СЛ требуют наличия достаточного количества обучающих примеров каждого класса атак, но не все виды атак в NSL-KDD представлены в достаточном для обучения количестве. Поэтому ИНС и СЛ были

обучены на основе примеров 9 атак, для которых в датасете было представлено по меньшей мере 500 образцов.

В рамках проведённого ранее исследования [28] ИНС и СЛ обучались на основе данных об атаках. Использованная в экспериментах в качестве классификатора ИНС представляла собой сеть прямого распространения (персептрон), скрытый слой которого содержит 16 нейронов с сигмоидной функцией активации, количество нейронов в выходном слое – 9 (по количеству распознаваемых классов, с функцией активации softmax). Коэффициент исключения (dropout) для регуляризации сети подобран экспериментально и равен 0,1.

Множество используемых образцов из NSL-KDD было разделено на обучающую, тестовую и контрольную выборки в соотношении 80 – 15 – 5. Контрольная выборка использовалась для предотвращения переобучения ИНС с реализацией раннего останова.

Параметры классификатора на основе СЛ подбирались с помощью процедуры поиска по сетке с перекрестной проверкой с тремя заходами – перебирались конкретные значения: количество деревьев, количество признаков, максимальная глубина и минимальное количество примеров в листовом узле. Матрицы ошибок ИНС и СЛ для тестовой выборки для каждого из 9-ти видов атак представлены на Рисунке 3.6.



а) искусственная нейронная сеть *б) случайный лес*

Рисунок 3.6 – Матрицы ошибок ИНС и СЛ для датасета NSL-KDD

Значения показателей эффективности классификаторов, полученные в результате эксперимента, сведены в Таблице 3.1, где Accuracy – доля верно классифицированных образов среди всех образов, рассматривается в разделе 2.2., вычисляется согласно (2.7); F_1 score – среднее гармоническое точности (Precision) и полноты (Recall), рассматривается в разделе 2.2., вычисляется согласно (2.8).

Таблица 3.1. – Значения показателей эффективности ИНС и СЛ в комбинации с ИИС

Классификаторы	Accuracy	F_1 score
ИИС + ИНС	0,997	0,997
ИИС + СЛ	0,999	0,999

Как видно из Таблицы 3.1, полученные результаты показывают эффективность объединения классификаторов в целом, а также демонстрируют, что объединение ИИС с алгоритмом случайного леса (СЛ) является более предпочтительным вариантом из рассмотренных. В данном случае ИИС оказывает содействие в определении того, является ли анализируемый образец соответствующим легитимному взаимодействию (отсутствие атаки) или нет, а СЛ отвечает за классификацию атаки. За счет этого уменьшаются ошибки первого и второго рода, и осуществляется достаточно точная классификация атак. Полученные результаты опубликованы в [28].

Но в данном случае возникает две проблемы:

1) СЛ в одиночку классифицирует атаки; если бы решение о классе атаки принималось с учетом мнения не только СЛ, но и других классификаторов, это бы уменьшило ошибки в определении класса атаки;

2) Если ИИС обнаружит неизвестную для СЛ атаку, то СЛ может некорректно ее классифицировать.

Один из вариантов решения – обучить ИИС классификации известных атак, но тогда могут возникать ситуации, когда ИИС сообщает об одном классе атаки, а СЛ – о другом. Возникает конфликт – проблема разрешения таких ситуаций.

Возможно, мнению одного из классификаторов стоит увеличить вес, но при этом допускаемые ошибки приоритетного алгоритма не будут чем-либо компенсированы. Необходим третий классификатор. Это увеличит точность классификации атаки, а также при параллельном независимом использовании трех классификаторов предположительно позволит нивелировать возможные ошибки ИИС в классификации «свой/чужой». Тестируемая система с использованием 3-х классификаторов была построена согласно схеме на Рисунке 3.7.

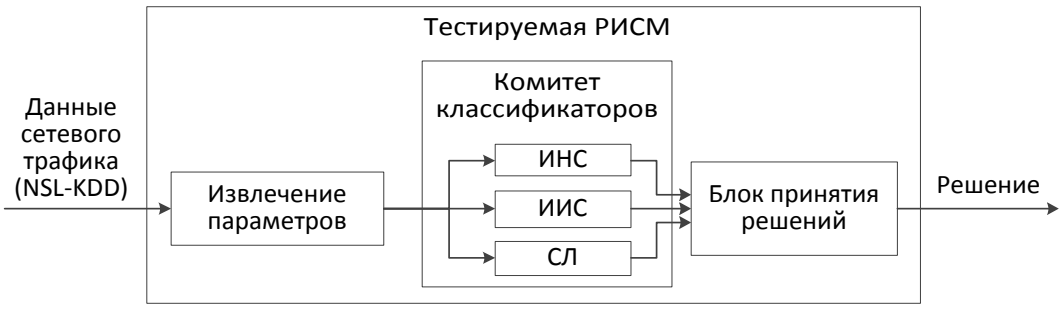


Рисунок 3.7 – Функционирование классификаторов в составе РИСМ в рамках второго эксперимента

Что касается механизма голосования, то, в первую очередь, была протестирована наиболее простая схема независимого голосования трех классификаторов по мажоритарному принципу, в соответствии с моделью простого объединения на уровне принятия решений. Взвешенное голосование не применялось, так как точность классификаторов была сопоставимо высокой.

Датасет NSL-KDD подвергся дополнительному балансированию с применением ресемплирования – SMOTE с KNN, для аугментации (расширения) маленьких классов до 5000 примеров и отбора по 15000 примеров из двух классов с большим количеством исходных данных.

Система была обучена заново. ИИС построена в варианте, предполагающем возможность классификации известных атак. Реализован алгоритм СЛ с оптимизацией гиперпараметров (перекрестная проверка, выбор по метрике

F_1 score). Параметры СЛ оценены на тестовой выборке, не участвовавшей в оптимизации гиперпараметров.

Созданная ИНС обучена на данных с контролем переобучения и ранним остановом. ИНС, аналогично СЛ, была проверена на основе тестовой выборки. Произведен выбор параметров архитектуры ИНС: количество нейронов в скрытом слое и коэффициент прореживания связей.

Установлено, что оптимальной является архитектура с 32 нейронами в скрытом слое, использование которой обеспечивает минимальный уровень ошибок. Полученные значения показателей эффективности использования ИНС и СЛ представлены в Таблице 3.2. Показатели данной таблицы: FNR, FPR и др. рассмотрены подробно в разд. 2.3.

Таблица 3.2 – Значения показателей эффективности ИНС, СЛ

Мера	ИНС	СЛ
FNR	0,003	0,001
FPR	0,013	0,003
TNR	0,987	0,997
TPR (Recall)	0,997	0,999
Precision	0,985	0,996
Accuracy	0,992	0,998
F_1score	0,991	0,997

Доля верно классифицированных атак каждого i-го класса рассчитывается следующим образом:

$$P_i = \frac{A_i}{A_{0,i}} \times 100\% \quad (3.4)$$

где $A_{0,i}$ – общее количество атак класса i; A_i – количество верно классифицированных атак класса i.

Значение данного показателя по каждому классификатору и каждому анализируемому виду атак представлено в Таблице 3.3, наилучшие значения по строке выделены желтым цветом.

Таблица 3.3 – Точность классификации по каждому виду атаки

Вид атаки	ИНС	СЛ	ИИС
back	0,9997	0,9997	0,9997
ipsweep	0,9942	0,9908	0,9914
neptune	0,9987	0,9996	0,9999
nmap	0,9792	0,9900	0,9926
portsweep	0,9942	0,9986	0,9997
satan	0,9659	0,9967	0,9950
Smurf	0,9996	0,9997	0,9966
teardrop	0,9997	0,9997	0,9997
warezclient	0,9921	0,9989	0,9685
Средняя взвешенная точность классификации атак	**0,9955**	**0,9984**	**0,9881**

Для наглядности данные Таблицы 3.8 представлены графически на Рисунке 3.7.

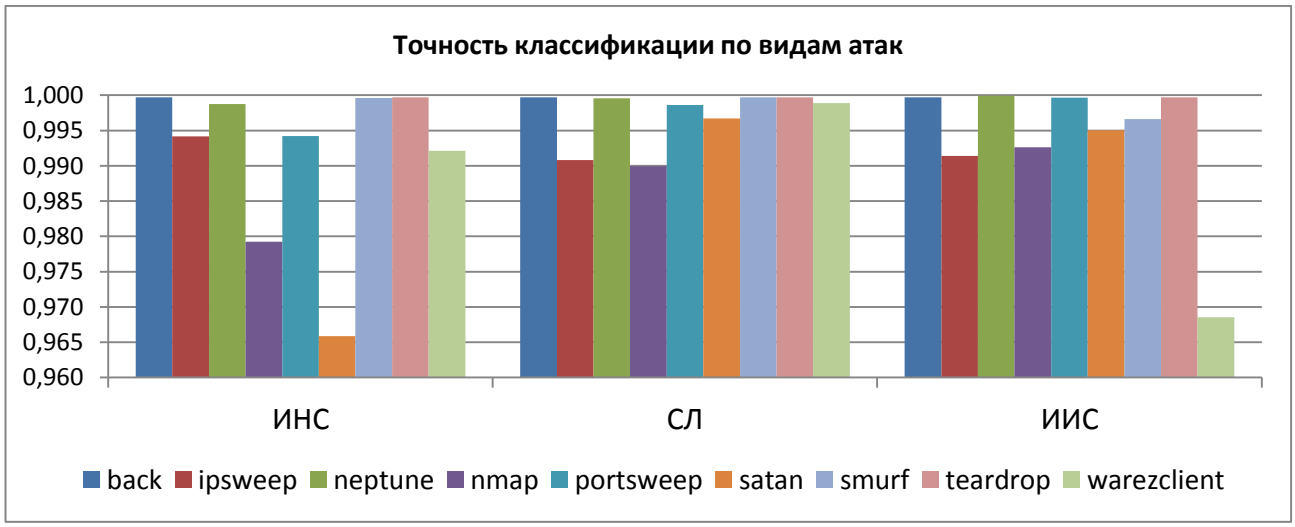


Рисунок 3.8 – Точность классификации по видам атак

Решение, принимаемое комитетом классификаторов, вырабатывалось подсистемой принятия решений (арбитром) по мажоритарному принципу.

Полученные результаты представлены в Таблице 3.4. Для наглядности лучшие значения также выделены цветом. Стоит отметить, что в данной таблице есть совпадения по лучшим значениям в строке, но одно из них выделено цветом в качестве лучшего, а другое – нет. Это связано с округлением: к примеру, FNR и для СЛ, и для комитета классификаторов в Таблице принимает одинаковое значение, равное 0,001. Однако представленные в Таблице значения округлены до тысячных. Если бы данные были представлены с округлением до десятитысячных, была бы заметна разница в пользу комитета классификаторов.

Таблица 3.4 – Значения показателей эффективности комитета классификаторов

Мера	ИНС	СЛ	ИИС	Среднее значение по ИНС, СЛ и ИИС	Комитет классификаторов
FNR	0,003	0,001	0,002	0,002	0,001
FPR	0,013	0,003	0,001	0,006	0,002
TNR	0,987	0,997	0,999	0,994	0,998
TPR (Recall)	0,997	0,999	0,998	0,998	0,999
Precision	0,985	0,996	0,999	0,993	0,998
Accuracy	0,992	0,998	0,999	0,996	0,998
F_1score	0,991	0,997	0,998	0,996	0,998

Таким образом, по четырем показателям из семи ИИС демонстрирует лучший результат. Проведенный анализ показал, что ИИС благодаря негативной селекции допускает меньше ошибок первого рода. В абсолютных значениях ИИС выдала ошибку первого рода (False Positive) по 25 проанализированным образцам трафика, в то время как СЛ совершил 101 такую ошибку, а ИНС – 136. Таким образом, возникали ситуации, когда ИИС верно определяла экземпляр легитимного трафика, а СЛ и ИНС одновременно ошибались, что при голосовании на основе большинства давало в результате ошибку и значения некоторых показателей эффективности комитета оказывались ниже, чем у ИИС. Поэтому механизм голосования был пересмотрен.

Рассматривался вариант увеличения веса мнения ИИС, но, с другой стороны, в точности классификации атак по 5 видам из 10 ИИС уступала другим

классификаторам. Поэтому было принято решение – организовать механизм голосования на основе двухуровневой схемы: на первом этапе рассматривается мнение ИИС, если она считает, что трафик соответствует нормальному (штатному) сетевому взаимодействию, то этого достаточно для определения экземпляра как соответствующего нормальному сетевому трафику. В противном случае, данные считаются соответствующими одной из атак и классифицируются на основе мнений большинства.

Если все три классификатора выдают три разных решения, приоритет отдается мнению СЛ, так как, согласно Таблице 3.3, его точность классификации атак выше. Полученные результаты при использовании двухуровневой схемы голосования представлены в Таблице 3.5, лучшие значения также выделены жёлтым цветом. Полученные результаты опубликованы в [29].

Таблица 3.5 – Значения показателей эффективности комитета классификаторов

Мера	ИНС	СЛ	ИИС	Среднее значение по ИНС, СЛ и ИИС	Комитет классификаторов
FNR	0,003	0,001	0,002	0,002	0,001
FPR	0,013	0,003	0,001	0,006	0,001
TNR	0,987	0,997	0,999	0,994	0,999
TPR (Recall)	0,997	0,999	0,998	0,998	0,999
Precision	0,985	0,996	0,999	0,993	0,999
Accuracy	0,992	0,998	0,999	0,996	0,999
F_1score	0,991	0,997	0,998	0,996	0,999

Полученное взвешенное значение точности классификации атак комитетом классификаторов представлено в Таблице 3.6, лучшее значение выделено желтым цветом.

Таблица 3.6 – Точность классификации атак

Показатель	ИНС	СЛ	ИИС	Комитет классификаторов
Доля верно классифицированных атак	0,9955	0,9984	0,9881	0,9988

Таким образом, при построении системы в соответствии со схемой, представленной на Рисунке 3.9, когда ИИС проводит первичный анализ, выделяет из всего потока данных образцы, соответствующие атакам, класс которых определяется в дальнейшем совместным решением ИИС, ИНС и СЛ, система демонстрирует более высокие значения показателей эффективности.

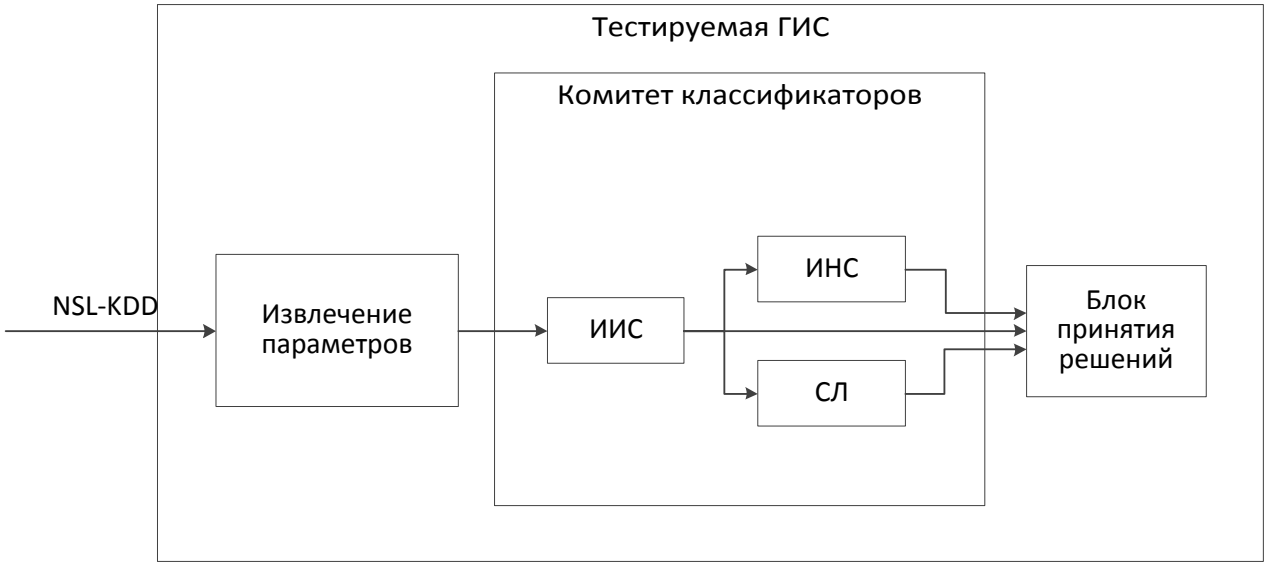


Рисунок 3.9 – Лучший рассмотренный вариант построения ГИС

Стоит отметить, что датасет NSL-KDD также использовался в [148] для построения IDS на основе ИИС в сравнении с такими известными системами обнаружения атак, как «Snort» [178] и «Bro» (сейчас система переименована в «Zeek») [199]. Число верно обнаруженных атак относительно всех проанализированных атак (Recall) для ИИС оказалось выше, чем у Snort примерно на 30%, чем у Bro – на 38%. Авторы аргументируют такое существенное отличие тем, что датасет NSL-KDD содержит более 90% данных об атаках в тестовой своей части, не содержащихся в обучающей части [148].

Была проведена другая серия вычислительных экспериментов по оценке эффективности разработанной ИИС отдельно (опубликовано в [30]) и в составе РИСМ на основе датасета Bot-IoT. В [108] авторами представлены показатели эффективности системы Snort с доработанными правилами для Bot-IoT.

Полученные значения показателей эффективности ИИС и РИСМ в сравнении с представленными в [108] показателями эффективности Snort приведены в Таблице 3.7. Показатель F_1 score для Snort был вычислен по (2.9) на основе значений Recall (Sensitivity) и Precision.

Таблица 3.7 – Показатели эффективности ИИС, РИСМ, Snort [108], полученные в результате анализа Bot-IoT

Классификатор	Recall (Sensitivity)	Precision	Accuracy	F_1 score
Snort	0,988	0,988	0,964	0,988
ИИС	0,994	0,996	0,995	0,995
РИСМ	0,999	0,999	0,999	0,999

Таким образом, только отдельный модуль разработанной ИИС, не требующий значительных вычислительных ресурсов, позволяет повысить значения показателей эффективности обнаружения атак на тестовом наборе данных в сравнении с существующими системами, но кроме того он обеспечивает возможность выявления новых, не известных атак. РИСМ демонстрирует повышение значений показателей эффективности на 1-3% по сравнению с существующими системами. Конечно, значения показателей эффективности обнаружения атак и аномалий в диапазоне 0,98-0,99 получены и справедливы для тестового набора данных, в реальных условиях, эти значения могут быть ниже.

Таким образом, разработан алгоритм функционирования РИСМ обнаружения атак и сетевых аномалий, основанный на первичном анализе данных посредством ИИС, выделении из всего потока только данных, соответствующих атакам и аномалиям, дальнейшая классификация которых осуществляется с помощью комитета классификаторов на основе ИИС, ИНС и СЛ и принятием решения по мажоритарной схеме.

Выводы по главе 3

1. Предложена концепция построения гибридной распределенной интеллектуальной системы (РИСМ) обнаружения атак и аномалий сетевого трафика промышленного Интернета вещей на основе объединения нескольких интеллектуальных систем – распределенной многоагентной двухуровневой ИИС, классификаторов на основе искусственной нейронной сети (ИНС) и алгоритма случайного леса (СЛ), подсистемы корреляционного анализа данных.

2. Интеграция в РИСМ подсистемы корреляционного анализа данных (КАД) позволяет дополнительно выделить наиболее критичные атаки, определить приоритеты в реагировании, скорректировать уровень опасности, используемый в двухуровневой ИИС на основе анализа данных, в том числе полученных от сторонних источников (SIEM-систем).

3. Разработан алгоритм функционирования РИСМ, основанный на первичном анализе данных сетевого трафика посредством двухуровневой ИИС, выделении из всего потока только тех данных, которые соответствуют атакам и аномалиям, их дальнейшей совместной классификации с использованием ИИС, ИНС и СЛ и принятием решения по мажоритарной схеме.

4 Разработка архитектуры интеллектуальной многоагентной системы мониторинга информационной безопасности промышленного Интернета вещей

4.1 Архитектура интеллектуальной многоагентной системы мониторинга информационной безопасности промышленного Интернета вещей. Методика обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей

Общие требования к источникам данных, сбору данных, их хранению, агрегированию и обработке, а также представлению результатов мониторинга ИБ и защите последних устанавливаются ГОСТ Р 59547-2021 «Защита информации. Мониторинг информационной безопасности. Общие положения» [4]. Структурный подход к реализации менеджмента инцидентов ИБ описан в ГОСТ Р ИСО/МЭК ТО 18044-2007 «Информационная технология. Методы и средства обеспечения безопасности. Менеджмент инцидентов информационной безопасности» [5], включая его цели и этапы их достижения: планирование и подготовку, использование, анализ и улучшение.

Разрабатываемая распределенная интеллектуальная система мониторинга (РИСМ) сетей IIoT предназначена для повышения уровня защищенности IIoT-систем от сетевых атак посредством выполнения следующих задач:

– сбор данных о сетевом взаимодействии IIoT-систем, выделении и нормализации параметров трафика;

– анализ параметров трафика распределенным комплексом агентов нижнего уровня двухуровневой искусственной иммунной системы (ИИС), выявление известных и неизвестных сетевых атак и аномалий;

– реализация граничных вычислений агентами верхнего уровня ИИС по оценке уровня опасности в каждой контрактной подсети IIoT для дифференцирования атак и аномалий, анализирующих в том числе информацию

от сторонних источников, таких как межсетевые экраны, системы мониторинга, SIEM, SCADA и пр.;

– классификация выявленных атак сервером с использованием комитета интеллектуальных систем, включая искусственную иммунную систему, искусственную нейронную сеть, алгоритм случайного леса;

– корреляционный анализ данных, позволяющий выделить наиболее критичные инциденты ИБ, определить приоритеты в реагировании.

Для построения интеллектуальной системы мониторинга ИБ IIoT используется многоагентный подход. Формальную модель многоагентной системы (МАС) можно представить в виде кортежа множеств:

$$\text{MAC} = < A, E, R, ORG, ACT, COM >, \quad (4.1)$$

где A – множество агентов; E – среда, в которой находится МАС; R –множество отношений (взаимодействий) между агентами; ORG – организационная структура МАС; ACT – планируемые стратегии поведения; COM – коммуникационные возможности.

С учетом данных вычислительных экспериментов, проведенных в разделе 3.3, разработана архитектура многоагентной РИСМ, включающая три уровня, представленная на Рисунке 4.1. Для наглядности уровни архитектуры РИСМ выделены зелеными пунктирными линиями, уровни архитектуры IIoT – тёмно-синими непрерывными линиями.

На нижнем уровне РИСМ функционирует расположенное в непосредственной близости к устройствам IIoT множество агентов, реализующих перехват или получение трафика, выявление атак и аномалий средствами ИИС, взаимодействующих друг с другом в части обмена обучающих данных, передающих информацию о выявленных атаках и аномалиях на второй уровень РИСМ. Агенты второго (среднего) уровня расположены на границах сетей IIoT, они получают данные об атаках от агентов нижнего уровня и (по возможности) от

подключаемых сторонних систем, таких как МЭ, SIEM, SCADA, выполняют оценку уровня опасности средствами ИИС, отделяют одиночные аномалии от реальных и потенциальных атак, передают данные на верхний уровень.

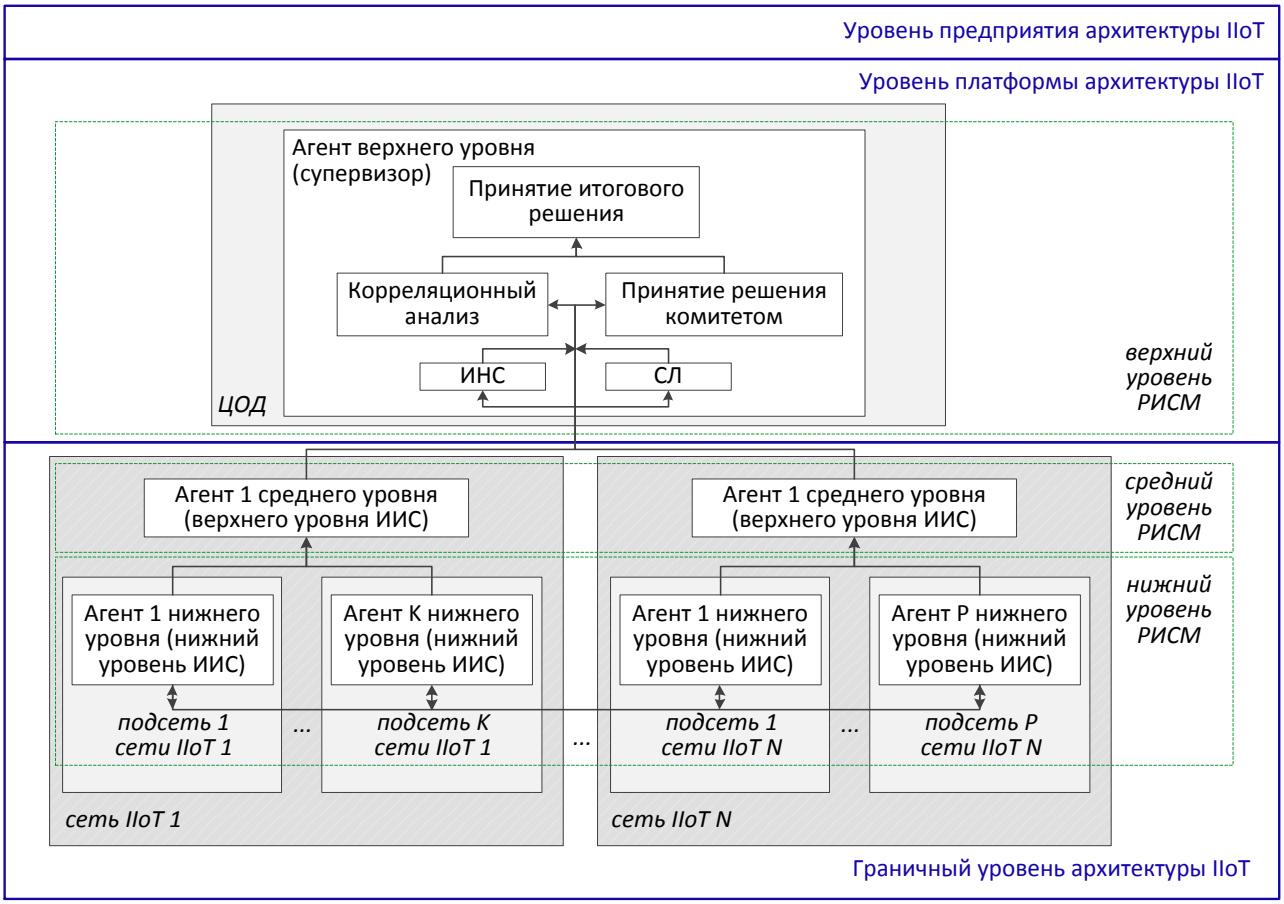


Рисунок 4.1 – Архитектура РИСМ ИБ IIoT

Агент верхнего уровня (супервизор) получает информацию о выявленных атаках, осуществляет дополнительный ее анализ подсистемами ИНС и СЛ, на основе мнений трёх интеллектуальных классификаторов (ИИС, ИНС, СЛ) принимает решение о классе атаки, проводит дополнительный корреляционный анализ данных (КАД), на основе решения комитета классификаторов и результата КАД принимает решение о критичности того или иного инцидента, передает информацию в консоль администрирования, осуществляет оповещение.

На этапе проведения исследований предполагается, что данные между агентами передаются в открытом виде по протоколам TCP/IP. В будущем

возможно обеспечение конфиденциальности и аутентичности взаимодействия агентов между собой криптографическими средствами, но данный вопрос в работе не рассматривается.

Таким образом, разработанная архитектура многоагентной распределенной интеллектуальной систем мониторинга ИБ промышленного Интернета вещей включает три уровня, первые два из которых функционируют на граничном уровне архитектуры IIoT, третий – на уровне платформы архитектуры IIoT.

Границы применимости системы. Агенты ИИС не требуют значительных вычислительных ресурсов, так как выполняют довольно простой с точки зрения вычислений алгоритм определения расстояния между анализируемыми векторами признаков данных и векторами-эталонами атак и аномалий на нижнем уровне ИИС и анализ сигналов опасности и определения уровня опасности на втором уровне двухуровневой ИИС, поэтому распределенная ИИС может быть внедрена в том числе в небольшие IIoT-системы, будучи реализованной на дополнительных устройствах, чтобы не нагружать существующие. Это обеспечит более высокий уровень эффективности (на 0,6-0,8% выше по показателям Recall, Precision, F_1 score и на 3,1% выше по показателю Accuracy) обнаружения существующих атак, а также позволит обнаруживать неизвестные атаки и аномалии, анализировать уровень опасности.

Если сеть IIoT достаточно большая, если у предприятия есть свой или арендованный центр обработки данных (ЦОД), или есть возможность использовать серверное оборудование, то технически целесообразно развертывание на предприятии РИСМ, нейронная сеть которой требует хорошего серверного оборудования, РИСМ обеспечивает больший прирост значений показателей эффективности (на 1,1% выше по показателям Recall, Precision, F_1 score, на 3,5% выше по показателю Accuracy) по сравнению с существующими системами.

Разработана методика обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей. Назначение данной методики – автоматизация процесса обнаружения атак и сетевых аномалий

промышленного Интернета вещей с использованием разработанной в работе гибридной системы мониторинга ИБ IIoT.

Входными данными для анализа являются параметры анализируемого сетевого трафика, выбранные с учетом особенностей объекта мониторинга (используемыми протоколами и пр.), получаемые посредством зеркалирования трафика или его перехвата.

Методика обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей заключается в последовательности следующих действий.

1. Определение используемых объектом мониторинга (IIoT) сетевых протоколов и технологий, выбор параметров сетевого взаимодействия для обнаружения атак и аномалий. Возможно использование параметров датасетов, выбранных способом, рассматриваемым в разделе 1.2.

2. Установка и распределение агентов нижнего уровня РИСМ по множеству подсетей IIoT, а на границах каждой сети – по одному агенту среднего уровня, на уровне платформы – агента верхнего уровня (супервизора).

3. Перехват посредством сниффинга или получение посредством зеркалирования данных о сетевом взаимодействии IIoT-устройств. Выделение и (или) расчет выбранных параметров сетевого трафика. Нормализация полученных данных.

4. Анализ данных средствами ИИС, реализуемыми агентами нижнего уровня, на предмет соответствия определенным видам атак или потенциальным неизвестным атакам, аномалиям.

5. Получение данных от других сторонних систем, таких как МЭ, SIEM, SCADA, определение уровня опасности на основе данных этих систем и данных предыдущего шага средствами ИИС, реализуемыми агентами второго уровня. Передача результатов супервизору.

7. Анализ супервизором полученных данных об атаках посредством ИНС и СЛ, КАД с учетом уровня опасности. Определение класса и критичности каждой атаки, оповещение администратора безопасности.

4.2 Решение прикладной задачи обнаружения компьютерных атак и аномалий сетевого трафика системы промышленного Интернета вещей с использованием РИСМ

Рассмотрим в качестве тестового объекта IIoT-систему контроля уровня и мутности воды в резервуаре (баке), которая, в свою очередь, входит в состав автоматизированной системы очистки и распределения воды в промышленных резервуарах.

С целью отработки алгоритмов мониторинга ИБ в данной системе был создан специальный испытательный стенд, с помощью которого решаются задачи сбора и анализа реальных данных промышленной сети Интернета вещей, включая задачу обнаружения реальных сетевых атак [201].

В состав программно-аппаратного обеспечения исследуемого объекта входят: программируемый логический контроллер (Programmable Logic Controller, PLC), системные журналы событий (логов), человеко-машинный интерфейс (Human Machine Interface, HMI), три сенсора и четыре актуатора (исполнительных механизма). В числе сенсоров: два датчика уровня (ДУ) воды в резервуаре и один аналоговый датчик мутности (ДМ). В число актуаторов, получающих команды от PLC, входят: трёхцветный световой индикатор, клапан, два водяных насоса. Для обеспечения возможности также ручного управления объект включает в себя кнопки включения, выключения, световой индикации. Для наглядности данные об элементах испытательного стенда [201] представлены в Таблице 4.1.

Таблица 4.1 – Элементы испытательного стенда

Тип элемента	Элемент	Количество
Сенсор	датчик уровня воды (Autonics CR18-8DP)	2
	датчик мутности (SEN0189)	1
Актуатор	световой индикатор	1
	клапан	1
	водяной насос (GA-2328ZZ)	2
Кнопка	кнопки включения, выключения и светодиодной индикации	3
PLC	PLC m241ce40	1

Продолжение таблицы 4.1

Тип элемента	Элемент	Количество
Сетевые технологии / протоколы	Ethernet	-
	Modbus	-
Средство сбора сетевого трафика	Argus [114]	1
	Wireshark [193]	1
Средство моделирования атак	Kali Linux Penetration Testing Distribution [140]	1
прочие элементы	HMI	1
	сервер логирования (logs)	1
	СОА	1
	МЭ	1
	Маршрутизатор	1

Схема взаимодействия элементов представлена Рисунком 4.2 [201].

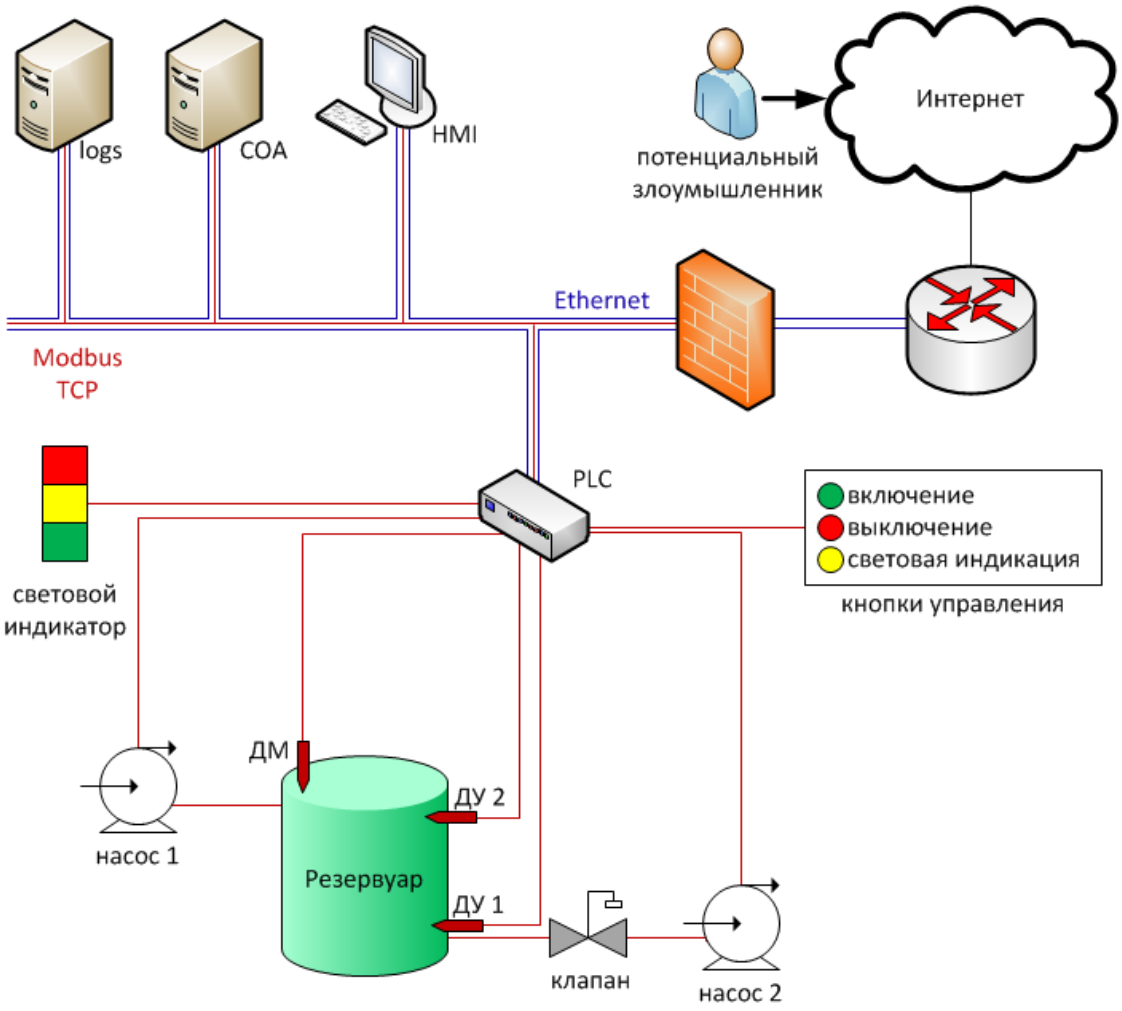


Рисунок 4.2 – Схема испытательного стенда [201]

Уровень воды в резервуаре должен быть в пределах диапазона, контролируемом двумя сенсорами. Превышение уровня выше допустимого максимума фиксирует ДУ2, он передает сигнал на PLC, PLC отключает насос 1, используемый для наполнения резервуара, открывает клапан, и включает насос 2, выкачивающий воду из бака. В противоположной ситуации, когда уровень воды в баке снижается ниже допустимого минимума, ДУ 1 фиксирует данный факт и передает информацию на PLC, который, в свою очередь, отключает насос 2, закрывает клапан, включает насос 1. Процесс повторяется, как только уровень воды превышает установленный максимум.

ДМ фиксирует мутность воды, передает данные на контроллер PLC, который оценивает уровень, используя два пороговых значения, включает свет индикатора соответствующего цвета: зеленый означает приемлемое значение мутности, красный – неприемлемое, желтый – значение находится между двумя пороговыми. Испытательный стенд получает данные сенсоров и статус системы от PLC и отображает через HMI [201].

Данные испытательного стенда были собраны его авторами с использованием Argus [114] и Wireshark [193], средняя скорость передачи данных составила 419 кбит/сек, средний размер пакета – 76,75 байт [200]. На основе собранных данных нормального сетевого взаимодействия и реализованных атак авторы стенда разработали и опубликовали датасет WUSTL-IIOT-2021 [194], предназначенный для исследований ИБ IIoT.

WUSTL-IIOT-2021 включает в себя сетевой трафик, характерный для нормального сетевого взаимодействия, а также для четырёх видов атак [200]:

– Backdoor (черный ход, задняя дверь). Атакующий использует предварительно установленное в систему, являющуюся объектом атаки, скрытое средство получения доступа для обхода легитимного процесса аутентификации с целью входа в систему, что позволяет получить доступ ко всем файлам и данным системы, запускать выполнение различных команд. В рассматриваемом случае объектом атаки является HMI-система, которая инфицируется вредоносным ПО Backdoor. Данное ПО работает скрытно от системного оператора SCADA,

открывает на HMI порт, позволяющий подключиться к нему удаленно. В результате атаки авторами испытательного стенда скачивается около гигабайта данных, создаются новые директории в системе SCADA и удаляются несколько файлов.

– Command Injection (внедрение команд). Атака заключается в несанкционированном внедрении зловредных команд в управляющую систему для нарушения нормального процесса её функционирования. Здесь объектом атаки является PLC. Атакующий подключается к сети для получения возможности чтения значений регистров PLC, затем перезаписывает одно из значений, важных для физического процесса. К примеру, когда насос 2 должен был выкачивать воду из резервуара, атакующий остановил его и запустил насос 1, и вода переполнила резервуар. Другой пример – подмена цвета индикации мутности воды.

– DoS (отказ в обслуживании). Выполняется атака на программный логический контроллер или человеко-машинный интерфейс посредством реализации большого потока ложных запросов синхронизации SYN или же HTTP-запросов к веб-серверу (GET или POST), с целью нарушения доступности системы. При успешной её реализации объект атаки становится загруженным ложными запросами и не может своевременно обрабатывать легитимные запросы, вследствие чего нарушается доступность системы.

– Reconnaissance (разведка). Атакующий подключается к сети и собирает информацию о системе, включая сведения о подключенных устройствах, используемых политиках безопасности, IP-адресах и так далее. После идентификации элементов сети атакующий строит схему сети для поиска уязвимостей. Атака чаще всего начинается с использованием снифферов – устройств, осуществляющих сбор и анализ проходящего через них сетевого трафика. Скрытое сканирование сети может проводиться между любыми узлами, к примеру, между человеко-машинным интерфейсом и программируемым логическим контроллером. Атака является пассивной и может не рассматриваться

в качестве серьезной, но сведения, полученные в результате её реализации, могут упростить проведение более серьезных атак.

Из параметров, содержащихся в датасете, авторами рекомендуется удалить параметры «StartTime», «LastTime», «SrcAddr», «DstAddr», «sIpId», «dIpId». Что является разумным, так как эти параметры характеризуют время регистрации события, IP-адреса, идентификаторы; это полезно для анализа самого события ИБ и идентификации оборудования, подвергнувшегося атаке, но не для обучения выявлению атак. Оставшиеся параметры представлены Таблицей 4.2 [194].

Таблица 4.2 – Параметры датасета WUSTL-IIOT-2021

Номер параметра	Параметр	Описание параметра
1	Mean flow (mean)	Средняя продолжительность активных потоков
2	Source Port (Sport)	Номер порта отправителя
3	Destination Port (Dport)	Номер порта назначения
4	Source Packets (Spkts)	Количество пакетов отправителя
5	Destination Packets (Dpkts)	Количество пакетов получателя
6	Total Packets (Tpkts)	Общее количество пакетов отправителя и получателя
7	Source Bytes (Sbytes)	Количество байт отправителя
8	Destination Bytes (Dbytes)	Количество байт получателя
9	Total Bytes (TBytes)	Общее количество байт отправителя и получателя
10	Source Load (Sload)	количество бит отправителя за секунду
11	Destination Load (Dload)	количество бит получателя за секунду
12	Total Load (Tload)	Общее количество бит за секунду суммарно отправителя и получателя
13	Source Rate (Srate)	Количество пакетов отправителя за секунду
14	Destination Rate (Drate)	Количество пакетов получателя за секунду
15	Total Rate (Trate)	Общее количество пакетов отправителя и получателя за секунду
16	Source Loss (Sloss)	Количество повторно переданных или отброшенных пакетов отправителя
17	Destination Loss (Dloss)	Количество повторно переданных или отброшенных пакетов получателя
18	Total Loss (Tloss)	Общее количество повторно переданных или отброшенных пакетов отправителя и получателя
19	Total Percent Loss (Ploss)	Процент повторно переданных или отброшенных пакетов (суммарно отправителя и получателя)
20	Source Jitter (ScrJitter)	Джиттер отправителя в миллисекундах
21	Destination Jitter (DrcJitter)	Джиттер получателя в миллисекундах
22	Source Interpacket (SIntPkt)	Межпакетный промежуток источника в миллисекундах

Продолжение Таблицы 4.2

Номер параметра	Параметр	Описание параметра
23	Destination Interpacket (DIntPkt)	Межпакетный промежуток получателя в миллисекундах
24	Protocol (Proto)	Протокол транзакции
25	Duration(Dur)	Общая продолжительность записи
26	TCP RTT (TcpRtt)	Время приема-передачи при установке TCP-соединения, сумма «synack» и «ackdat».
27	Idle Time (Idle)	время с момента последней пакетной активности.
28	Sum (sum)	общая накопленная продолжительность агрегированных записей
29	Min (min)	минимальная продолжительность агрегированных записей
30	Max (max)	максимальная продолжительность агрегированных записей
31	Source Diff Serve Byte (sDSb)	отличающееся значение serve byte отправителя
32	Source TTL (sTtl)	Значение TTL от отправителя к получателю
33	Destination TTL (dTtl)	Значение TTL от получателя к отправителю
34	Source App Byte (SAppBytes)	Исходящие байты приложения
35	Destination App Byte (DAppBytes)	Входящие байты приложения
36	Total App Byte (TotAppByte)	Общее количество байт приложения (входящих и исходящих)
37	SYN_Ack (SynAck)	Время установки TCP-соединения, время между пакетами SYN и SYN ACK
38	Run Time (RunTime)	общее время работы активного потока. Это значение создается путем агрегирования и представляет собой сумму продолжительности записей.
39	Source TOC (sTos)	Значение байта «тип сервиса» отправителя
40	Source Jitter (SrcJitAct)	Неактивный (idle) джиттер источника в миллисекундах
41	Destination Jitter (DstJitAct)	Активный джиттер получателя в миллисекундах
42	Traffic	Соответствие нормальному трафику или конкретному виду атаки
43	Target	Флаг соответствия атаки. 1 – атака, 0 – норма.

Датасет был нормализован рассмотренным в разд. 2.3 способом в различных вариациях диапазонов параметров. В первой серии экспериментов был выбран диапазон [0; 255]. Подход позволил сжать пространство параметров до 9. Эффективность сжатия оценивалась по показателю P:

$$P_n = \frac{|A_m|}{|A|} \quad (4.2)$$

где A – множество атак; $|A|$ – размерность множества A; A_m – множество атак, имеющих идентичные элементы во множестве нормальных соединений N, то есть $A_m = A \cap N$; $|A_m|$ – размерность множества A_m; n – число параметров.

Если P равен нулю, это значит, что множества A и N не пересекаются вообще; если показатель P имеет низкое значение, значит, существует небольшое количество точек в пространстве параметров, соответствующих и атакам, и нормальному состоянию, которые, в зависимости от приоритета обучения, обеспечат непреодолимый уровень ошибок первого или второго рода. Нормализованные параметры ранжированы по наименьшему проценту совпадений.

По 15 параметрам значение P_{15} оказалось примерно равным $6{,}4 \times 10^{-4}$. Довольно низкое значение, но анализ совпадающих по этим 15 параметрам строк показал, что они отличаются только по совокупности значений параметров SIntPkt, SrcRate, Mean; их номера после ранжирования были 7, 20 и 21 соответственно. Вместо увеличения количества параметров до 21, был изменен порядок параметров и номера для SrcRate, Mean, определён как 8 и 9, увеличено число анализируемых параметров до 17 и значение P резко упало. Количество параметров постепенно уменьшалось, пока P оставалось на достаточно низком уровне. Полученный порядок приоритета параметров представлен в Таблице 4.3, курсивом выделены SrcRate и Mean с измененными порядковыми номерами.

Таблица 4.3 – Ранжированные параметры WUSTL-IIOT-2021

Номер параметра после ранжирования	Начальный номер параметра в датасете	Сокращенное наименованием параметра	Процент совпадений по параметру
1	3	Dport	0,41%
2	32	sTtl	0,58%
3	2	Sport	10,87%
4	19	pLoss	46,32%

Продолжение Таблицы 4.3

Номер параметра после ранжирования	Начальный номер параметра в датасете	Сокращенное наименованием параметра	Процент совпадений по параметру
5	40	SrcJitAct	49,80%
6	27	IdleTime	56,64%
7	22	SIntPkt	60,53%
8	13	*SrcRate*	89,61%
9	1	*Mean*	91,99%
10	10	SrcLoad	69,05%
11	12	Load	69,95%
12	20	SrcJitter	77,22%
13	34	SAppBytes	77,23%
14	24	Proto	79,45%
15	9	TotBytes	80,84%
16	8	SrcBytes	80,85%
17	38	RunTime	81,88%
18	30	Max	82,21%
19	29	Min	82,22%
20	28	Sum	82,22%
21	15	Rate	89,06%
22	25	Dur	92,08%
23	36	TotAppByte	96,18%
24	18	Loss	97,20%
25	16	SrcLoss	97,22%
26	14	DstRate	97,28%
27	17	DstLoss	97,74%
28	26	TcpRtt	98,04%
29	35	DAppBytes	98,64%
30	11	DstLoad	98,88%
31	7	DstBytes	99,62%
32	5	DstPkts	99,63%
33	6	TotPkts	99,65%
34	4	SrcPkts	99,65%
35	23	DIntPkt	99,83%
36	21	DstJitter	99,84%
37	41	DstJitAct	99,84%
38	33	dTtl	99,88%
39	39	sTos	99,93%
40	31	sDSb	99,94%
41	37	SynAck	99,96%

Для девяти параметров был получен результат $P_9 \approx 6{,}7 \times 10^{-5}$, что является вполне приемлемым и не сказывается существенно на ошибках

используемых интеллектуальных классификаторов. Увеличение количества используемых параметров для снижения P не оправдано, так как это снизит производительность. Уменьшение до 8 – увеличивает P на порядок. Стоит отметить, что по 9 параметрам множества различных классов атак не пересекались вообще.

Таким образом, для случая применения WUSTL-IIOT-2021, где приемлемы значения параметров в диапазоне [0;255], целесообразно применение первых 9 ранжированных параметров, представленных в Таблице 4.3. Вместе с тем, применение в нашей системе ИНС предполагает нормализацию ее параметров в диапазоне [0; 1]. Если рассматривать округление до сотых, то диапазон значений каждого параметра равен 100. Тогда целесообразно сразу нормализовать данные в диапазоне [0; 100], что и было выполнено. Получен результат:

– по 40 параметрам: $P_{40} \approx 3{,}45 \times 10^{-4}$;

– по 14 параметрам: $P_{14} \approx 3{,}68 \times 10^{-4}$;

– по 13 параметрам: $P_{13} \approx 8{,}39 \times 10^{-4}$.

Рационально применение 14 параметров, увеличение их числа не дает значительного снижения P, а уменьшение – приводит к росту P в разы. Множества различных классов атак по 14 параметрам не пересекаются совсем.

WUSTL-IIOT-2021 также был нормализован в диапазоне [0; 64], но в таком случае хоть и достигались приемлемые значения P, но возникали пересечения между множествами различных классов атак, которые представлены в малом количестве. 72 строки были общими для множеств Backdoor и Reconnaissance, а это 34% и 28% соответственно, недопустимые уровни.

Таким образом, получено два приемлемых варианта. Это использование:

1) 9 параметров в диапазоне [0; 255] при $P_9 \approx 6{,}7 \times 10^{-5}$;

1) 14 параметров в диапазоне [0; 100] при $P_{14} \approx 3{,}68 \times 10^{-4}$.

Верхний уровень системы мониторинга ИБ (супервизор) является централизованным, расположен на средстве вычислительной техники с

достаточным уровнем ресурсов, здесь можно использовать оба варианта, а так как используется ИНС, предпочтительнее первый вариант.

Вопрос обеспечения высоких показателей быстродействия ИИС, учитывая, что она распределена по подсетям и может быть реализована на маломощном оборудовании, является быть критичным. Увеличение диапазона значений каждого параметра потенциально негативно влияет на быстродействие, так же как и увеличение количества параметров. Поэтому необходимо проведение вычислительных экспериментов ИИС в двух этих вариантах. Вычислительные эксперименты проводились и по первому варианту, и по второму одновременно, каждый в отдельном потоке на одном компьютере с процессором AMD Athlon™ X4 870K Quad Core 3.9 ГГЦ. Сначала использовалось пороговое расстояние Хэмминга, равное двум ($Threshold = 2$). То есть, если значения двух параметров у сравниваемых векторов отличаются, вектора считаются значительно отличающимися. Было сгенерировано три типа детекторов:

– D_1 – для обнаружения и классификации известных атак, обучены на половине данных о нормальном сетевом взаимодействии ($|N|/2$) и половине данных об атаках ($|A|/2$), количество D_1 получилось меньше количества строк ($|A|/2$), так как некоторые детекторы не прошли негативную селекцию и были отброшены;

– D_2 – для дополнения и одновременного применения с D_1 в целях обеспечения возможности обнаружения неизвестных атак, обучались только на ($|N|/2$) с обеспечением удаленности от D_1;

– D_3 – для тестирования принципиальной способности обнаружения неизвестных атак, обучались только на ($|N|/2$) для применения в качестве единственного вида детекторов и выявления атак датасета, которые условно являются незнакомыми системе.

Полученные результаты представлены в Таблице 4.4, жёлтым цветом выделены лучшие значения.

Таблица 4.4 – Полученные результаты (*Threshold* = 2; два вида детекторов)

Характеристика	Вариант 1	Вариант 2
Количество параметров	9	14
Диапазон значений параметров	[0; 255]	[0; 100]
Пороговое расстояние Хэмминга	2	
Количество детекторов D_1, выявляющих известные атаки	43 460	
Количество детекторов D_2, выявляющих неизвестные атаки	456 540	
Общее количество детекторов	500 000	
Продолжительность формирования детекторов D_1	0:39:52	0:39:38
Продолжительность формирования детекторов D_2	8:30:36	8:28:59
Общее время формирования детекторов	9:10:28	9:08:37
Время анализа половины данных об атаках (43 508 образцов)	0:04:37	0:03:34
Доля ошибок второго рода (False Negatives)	0,138	0,089
Доля ошибок первого рода (False Positives)	<0,001	<0,001

Результаты демонстрируют сопоставимое время обучения и анализа, в первом варианте несколько быстрее осуществляется анализ, во втором – формирование детекторов, но также ниже уровень ошибок второго рода (FNR). Полученные значения FNR являются неудовлетворительными.

Также проведена другая серия экспериментов по формированию только детекторов D_3 для оценки способности и времени на обнаружение неизвестных атак, результаты представлены в Таблице 4.5, жёлтым цветом выделены лучшие значения.

Таблица 4.5 – Полученные результаты (*Threshold* = 2; один вид детекторов)

Характеристика	Вариант 1	Вариант 2
Количество параметров	9	14
Диапазон значений параметров	[0; 255]	[0; 100]
Пороговое расстояние Хэмминга	2	
Количество детекторов D_3, выявляющих неизвестные атаки	500 000	
Продолжительность формирования детекторов D_3	6:52:36	6:50:44
Время анализа половины данных об атаках (43 508 образцов)	0:04:37	0:03:34
Доля ошибок второго рода (False Negatives)	1,000	1,000
Доля ошибок первого рода (False Positives)	<0,001	<0,001

Такой подход привел к уровню FNR, равному единице, то есть система не смогла обнаружить ни одной атаки, значит, пространство атак охватывается детекторами недостаточно, для следующей серии экспериментов пороговое

расстояние увеличено до 3, результаты представлены в Таблице 4.6, жёлтым цветом выделены лучшие значения.

Таблица 4.6 – Полученные результаты (*Threshold* = 3; два вида детекторов)

Характеристика	Вариант 1	Вариант 2
Количество параметров	9	14
Диапазон значений параметров	[0; 255]	[0; 100]
Пороговое расстояние Хэмминга	3	
Количество детекторов D_1, выявляющих известные атаки	60 338	
Количество детекторов D_2, выявляющих неизвестные атаки	439 662	
Общее количество детекторов	500 000	
Продолжительность формирования детекторов D_1	00:57:53	00:58:03
Продолжительность формирования детекторов D_2	7:44:29	7:45:55
Общее время формирования детекторов	8:42:22	8:43:58
Время анализа 30% атак (26 104 образцов)	00:00:43	00:00:45
Время анализа 50% нормы (553 724 образцов)	1:16:37	1:15:27
Доля ошибок второго рода (False Negatives)	0,017	0,013
Доля ошибок первого рода (False Positives)	<0,001	<0,001

В данном и последующих экспериментах число атак, на основе которых обучались D_1, увеличено до $\left(\frac{2}{3}|A|\right)$, число атак, на которых оценивалась эффективность, соответственно, уменьшено до $\left(\frac{1}{3}|A|\right)$, время обучения оказалось лучшим уже для первого варианта по сравнению со вторым. По времени анализа нет однозначного лидера. Таким образом, по быстродействию оба варианта можно считать примерно эквивалентными. Полученные значения FNR значительно снизились, но остаются выше допустимого уровня.

Анализ неизвестных атак со значением *Threshold* = 3 аналогично значению *Threshold* = 2 показал уровень FNR = 1, далее также повышалась вероятность обнаружения атак увеличением области, покрываемой каждым детектором за счет увеличения *Threshold*. Каждый раз детекторы генерировались заново, неизвестные атаки стали обнаруживаться для первого варианта при *Threshold* = 5, для второго при *Threshold* = 9 и продолжали обнаруживаться даже при количестве детекторов уменьшенном с 500 000 до 50 000.

Сначала уровень ошибок FNR был высоким, но ИИС дообучается в процессе анализа. В первом варианте для достижения значения FNR ≈ 0,004 потребовалось 64 эпох. Во втором варианте за 40 эпох было достигнуто значение FNR < 0,001, но возник недопустимый уровень ошибок первого рода (FPR ≈ 0,3).

Полученные результаты представлены в Таблице 4.7, красным цветом выделены неприемлемые значения, зеленым – приемлемые. Таким образом, приемлемым оказалось только применение первого варианта построения ИИС, при котором анализируется 9 параметров в диапазоне значений [0; 255], при значении *Threshold* = 5.

Таблица 4.7 – Уровни ошибок FNR и FPR в зависимости от *Threshold*.

№ варианта	Значение *Threshold*	FNR по известным атакам (D_1)	FPR по известным атакам (D_1)	FNR по неизвестным атакам (D_3)	FPR по неизвестным атакам (D_3)
1	3	0,012	< 0,001	> 0,999	< 0,001
	4	0,002	< 0,001	> 0,999	< 0,001
	5	0,005	< 0,001	0,004	< 0,001
	6	0,099	< 0,001	< 0,001	0,543
2	4	0,011	< 0,001	> 0,999	< 0,001
	5	0,008	< 0,001	> 0,999	< 0,001
	8	0,022	< 0,001	> 0,999	< 0,001
	9	0,101	< 0,001	< 0,001	0,300

Если с анализом известных атак всё очевидно: система обучена на одной части данных, протестирована на другой, и сразу демонстрирует высокие значения показателей эффективности обнаружения, то анализ неизвестных атак следует рассмотреть несколько подробнее. В первую очередь, система обучена на части данных о нормальном сетевом взаимодействии, затем ей продемонстрирована треть данных об атаках. При первой эпохе анализа система способна выявить всего несколько атак из более чем 26 тысяч, но уже при следующей эпохе анализа, эффективность увеличивается и достигает 0,996 к 64-й эпохе. Динамика изменения TPR (Recall) представлена Рисунком 4.3.

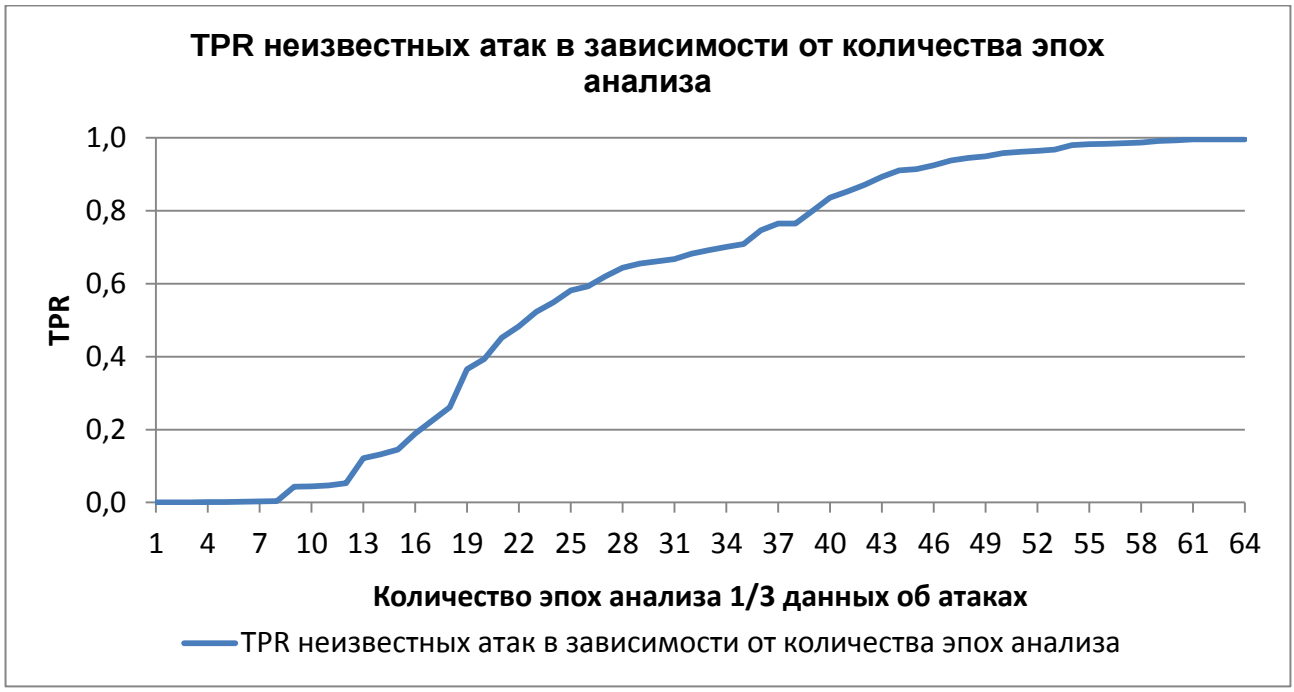


Рисунок 4.3 – Динамика изменения TPR (Recall) для неизвестных атак

Такой подход осуществляет перекрытие не всего пространства аномалий, но достаточной его части, динамически изменяемой случайным образом. При данном подходе достаточно выявления некоторых экземпляров атак, после чего детекторами осуществляется более плотное перекрытие близкого к ним пространства. Это позволяет использовать меньшее количество детекторов, что улучшает производительность и позволяет избегать высокого уровня ошибок первого рода (FPR), который зачастую приводит к игнорированию операторами или администраторами всех сигналов тревог (алертов) системы.

В случае необходимости параноидального обнаружения каждой аномалии можно рассмотреть обратный подход, где детекторы D_N обнаруживают не атаки, а образы (паттерны) нормального взаимодействия. Если ни один детектор D_N не соответствует паттерну, он отмечается как атака. Рассмотрено формирование таких детекторов с использованием двух третей данных N для обучения. Для тестирования использовались оставшаяся треть данных N и полный набор A. На первой же эпохе анализа при *threshold* = 5 выявлено 99,06% атак, то есть TPR ≈ 0,990, FNR ≈ 0,010, TNR ≈ 0,999, FPR ≈ 0,001. При уменьшении *threshold*

растут ошибки первого рода, а ошибки второго рода не снижаются. При увеличении *threshold* растут ошибки второго рода.

Подход формирования D_N для обнаружения неизвестных атак сразу позволяет получить высокое значение одного из ключевых показателей эффективности – Recall, но оно получается ограниченным значением 0,990, а формирование D_3 позволяет хоть и не сразу, но выявить 0,996 и на этом не останавливается, стремится к 1 с каждой эпохой анализа. Поэтому рекомендуется построение детекторов D_1 для высокоэффективного выявления известных атак сразу же, и дополнения детекторами D_2 для обеспечения возможности выявления неизвестных атак. Что и было реализовано для обнаружения сетевых атак на рассматриваемый объект.

Были построены классификаторы на основе алгоритма случайного леса (СЛ) и искусственной нейронной сети (ИНС). СЛ и ИНС для лучшей точности классификации анализировали полный состав параметров, нормализованный в диапазоне [0; 100]. Таким образом, подсистема перехвата или получения трафика должна выделять полный набор из 41 параметра, 9 из них приводить к диапазону [0; 255], если данные соответствуют атаке, то агентами нижнего уровня весь набор из 41 параметра соответствующих данных приводится к диапазону [0; 100] для дальнейшего их анализа посредством СЛ и ИНС.

Для построения классификаторов на основе СЛ и ИНС датасет разделен на выборки: обучающую, тестовую и проверочную в соотношении 60% – 25% – 15%. Архитектура ИНС включает:

- входной слой;
- скрытый слой с функцией активации – сигмоида, 15 нейронов;
- выходной слой, 5 нейронов, функция активации softmax.

Алгоритм оптимизации весовых коэффициентов – Adamar, скорость обучения – адаптивная, количество эпох обучения – 50, контроль раннего останова на проверочном множестве. Проведена оптимизация гиперпараметров на

сетке (перекрестная проверка с 10 заходами, перебор 100 моделей), результат представлен Рисунком 4.4.

```
Количество нейронов в скрытом слое = [9, 10, 12, 14, 16]
Коэффициент контрастирования весов = [1, 2, 3, 4, 5]
Коэффициент dropout = [0.0, 0.1, 0.2, 0.3]
Наилучшая модель: {'dropout_rate': 0.0, 'neurons': 9, 'weight_constraint': 1
```

Рисунок 4.4 – Оптимизация гиперпараметров ИНС

Оптимизация классификатора Случайный лес представлена Рисунком 4.5.

```
        Количество деревьев: [30, 50, 70],
        Минимальное количество в листе: [3, 5, 7],
        Максимальное количество признаков для разделения: [3, 5, 7, 10],
        Максимальная высота дерева: [10, 17, 23]}

Наилучшая модель:
        {'max_depth': 23,
         'max_features': 10,
         'min_samples_leaf': 3,
         'n_estimators': 30}
```

Рисунок 4.5 – Оптимизация классификатора Случайный лес

На Рисунке 4.6 представлена оценка F_1 score и точности классификации для обучающей и тестовой выборок в зависимости от максимальной высоты дерева.

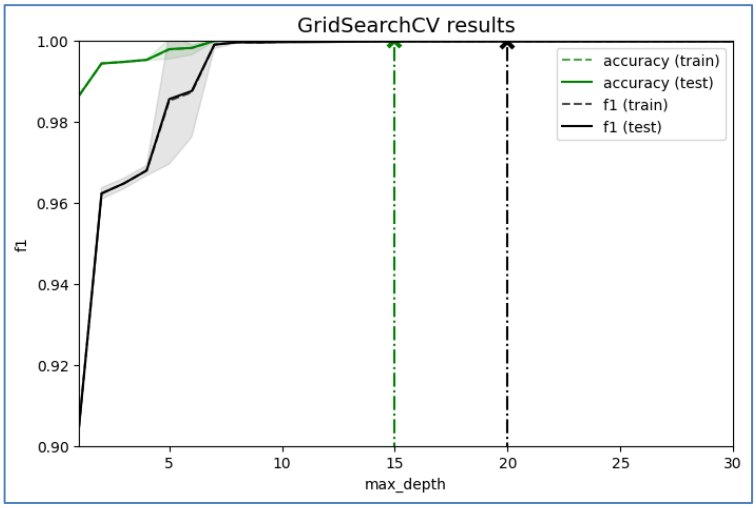
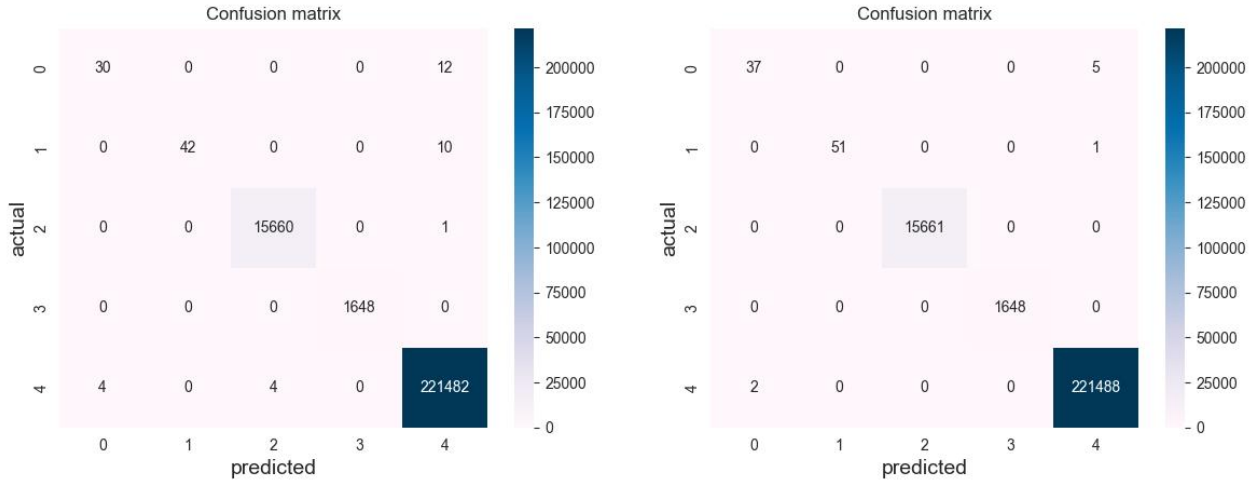


Рисунок 4.6 – Оценка F_1 score и точности классификации для обучающей и тестовой выборок в зависимости от максимальной высоты дерева

Матрицы ошибок СЛ и ИНС представлены на Рисунке 4.7, где классы расположены слева направо и сверху вниз в порядке: 0 – Backdoor, 1 – Command Injection (Comm), 2 – DoS, 3 – Reconnaissance (Reconn), 4 – Normal. Подборка архитектуры СЛ заняла 8 часов, построение одиночного СЛ с подобранными параметрами – 15 минут, одиночной ИНС – 150 минут.



а) искусственная нейронная сеть *б) случайный лес*

Рисунок 4.7 – Матрицы ошибок СЛ и ИНС на основе WUSTL-IIOT-2021

Результаты классификации «свой/чужой» представлены в Таблице 4.8, лучшие значения выделены жёлтым цветом.

Таблица 4.8 – Значения показателей эффективности классификации «свой/чужой»

Показатель	СЛ	ИНС	ИИС ($D_1 + D_2$)	Комитет при равнозначном голосовании	Комитет при приоритете ИИС
TN	221 488	221 482	221 488	221 488	221 488
TP	17 397	17 380	17 400	17 398	17 400
FP	2	8	2	2	2
FN	6	23	3	5	3
TNR	0,99999	0,99996	0,99999	0,99999	0,99999
TPR	0,99965	0,99868	0,99983	0,99971	0,99983
FPR	0,00001	0,00004	0,00001	0,00001	0,00001
FNR	0,00034	0,00132	0,00017	0,00028	0,00017
Precision	0,99989	0,99954	0,99989	0,99989	0,99989
Recall	0,99966	0,99868	0,99983	0,99971	0,99983

Продолжение Таблицы 4.8

Показатель	СЛ	ИНС	ИИС ($D_1 + D_2$)	Комитет при равнозначном голосовании	Комитет при приоритете ИИС
Accuracy	0,99997	0,99987	0,99998	0,99997	0,99998
F_1 score	0,99977	0,99911	0,99986	0,99980	0,99986

Классификаторы допускают небольшое количество ошибок, но по всем показателям в разделении множеств атак и нормального состояния лидирует ИИС. При построении равнозначного голосования получены результаты хуже, чем для ИИС, поэтому приоритет в классификации «свой/чужой» передан ИИС.

Однако в классификации известных атак ИИС существенно проигрывает СЛ и ИНС, возникали не только ошибки присвоения известного класса атаки, но и часть атак была классифицирована как «unknown». Ошибки определения класса атак сведены в Таблице 4.9, наименьшие значения ошибок по строке выделены желтым цветом.

Таблица 4.9 – Ошибки определения класса атак

Класс атаки	Кол-во анализи-руемых атак	Количество ошибочно классифицированных атак в абсолютных (абс.) единицах и относительно (относ.) всего количества примеров конкретного вида атак							
		СЛ		ИНС		ИИС		Комитет при равнозначном голосовании	
		абс.	относ.	абс.	относ.	абс.	относ.	абс.	относ.
DoS	15 661	0	0,0000	1	0,0001	1	0,0001	0	0,0000
Reconn	1 648	0	0,0000	0	0,0000	59	0,0358	0	0,0000
Comm	52	1	0,0192	10	0,1923	26	0,5000	1	0,0192
Backdoor	42	5	0,1190	12	0,2857	8	0,1905	4	0,0952
ИТОГО	17 403	6	0,0003	23	0,0013	94	0,0054	5	0,0003

По данным Таблицы 4.9, при равнозначном голосовании на основе мнения большинства лучшую точность определения класса атаки показывает комитет классификаторов.

Показатели эффективности СЛ были сравнены с показателями эффективности других алгоритмов машинного обучения, результаты

представлены в Таблице 4.10, где лучшие значения показателей эффективности по каждому столбцу выделены желтым цветом, дополнительно использованы следующие метрики:

$$MCC = \frac{(TP \times TN - FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}; \quad (4.3)$$

$$AUC = \frac{TNR+TPR}{2}; \quad (4.4)$$

$$Kappa = \frac{(TP+FP)(TP+FN)+(FN+TN)(FP+TN)}{(TP+FP+FN+TN)^2}. \quad (4.5)$$

Таблица 4.10 – Сравнение результатов обучения классификаторов

Классификатор	Accuracy	AUC	Recall	Precision	F_1 score	Kappa	MCC	Время построения
дерево решений	1,0000	0,9999	1,0000	1,0000	1,0000	0,9997	0,9997	0:01:49
СЛ	1,0000	1,0000	1,0000	1,0000	1,0000	0,9998	0,9998	0:15:02
экстра-деревья	1,0000	0,9999	1,0000	1,0000	1,0000	0,9997	0,9997	0:13:16
экстремальный градиентный бустинг (ГБ)	1,0000	1,0000	1,0000	1,0000	1,0000	0,9998	0,9998	2:11:20
ГБ с CatBoost	1,0000	1,0000	1,0000	1,0000	1,0000	0,9998	0,9998	3:17:15
k ближайших соседей	0,9999	0,9999	0,9999	0,9990	0,9999	0,9993	0,9993	2:10:38
Логистическая регрессия	0,9994	0,9999	0,9994	0,9994	0,9994	0,9953	0,9953	0:44:25
SVM	0,9988	0,0000	0,9988	0,9988	0,9988	0,9910	0,9911	0:07:01
ГБ	0,9988	0,9983	0,9988	0,9988	0,9993	0,9917	0,9919	4:36:46
квадратичный дискриминантный анализ	0,9919	0,9984	0,9919	0,9928	0,9921	0,9433	0,9448	0:03:55
линейный дискриминантный анализ	0,9914	0,9990	0,9914	0,9926	0,9917	0,9396	0,9409	0:04:12
гребневый классификатор	0,9913	0,0000	0,9913	0,9919	0,9911	0,9376	0,9380	0:01:07
NBK	0,9848	0,9962	0,9848	0,9925	0,9983	0,8984	0,9030	0:01:18
машина легкого ГБ	0,9354	0,8426	0,9354	0,9535	0,9419	0,6148	0,6293	0:07:27
Dummy-классификатор	0,9272	0,5000	0,9272	0,5960	0,8921	0,0000	0,0000	0:00:46
Адаптивный бустинг (Ada)	0,8632	0,9527	0,8632	0,9563	0,8943	0,5151	0,5811	0:18:20

По данным Таблицы 4.10 видно, что лучшие результаты по рассмотренным показателям эффективности, не учитывая показатель затрат времени, показывают СЛ, экстремальный градиентный бустинг (ГБ) и ГБ с CatBoost. Наилучший показатель по затратам времени из указанных трёх классификаторов принадлежит алгоритму случайного леса. Стоит отметить, что здесь оценивалось время построения классификаторов с уже подобранными параметрами. Также данные Таблиц 4.10 и 4.8 отличаются для СЛ, так как Таблица 4.8 содержит результаты анализа тестовой части датасета, а Таблица 4.10 – обучающей.

Таким образом, результаты вычислительных экспериментов демонстрируют высокую эффективность предлагаемой гибридной распределенной интеллектуальной системы мониторинга информационной безопасности промышленного Интернета вещей для обнаружения и классификации известных сетевых атак, способности к обнаружению неизвестных, обучению на их основе выявлению других неизвестных атак схожих с обнаруженными.

4.3 Интеллектуальная система обнаружения аномалий сетевого трафика беспроводной сенсорной сети IIoT

Беспроводные сенсорные сети (WSN) представляют интерес, так как являются одной из главных системообразующих компонент IIoT. Для обучения и тестирования систем обнаружения атак на WSN в [107] построена модель WSN, смоделированы сетевые атаки, собран и размечен сетевой трафик в составе датасета WSN-DS [106], используемого в данной работе с непосредственного одобрения его авторов.

Согласно [107], сетевое взаимодействие данной модели построено на основе протокола LEACH, для моделирования атак использовался инструмент «Network Simulator NS2» [181]. Все эксперименты проводились с использованием компьютера Intel® Core™ i3 CPU M 380, 2.53 ГГц, 4 ГБ ОЗУ, со следующими условиями:

– сенсорные узлы статичны;

– базовые станции (BS) не имеют ограничений в энергии;

– вредоносные узлы имеют большую дальность передачи, чем нормальные узлы;

– у каждого сенсорного узла есть данные для передачи на каждом интервале времени.

Модель включает 100 узлов, расположенных случайным образом на квадратной поверхности 100 м × 100 м. Каждый сенсор принимает участие в мониторинге 5 соседей [106]. Параметры модели сведены в Таблице 4.11 [107].

Таблица 4.11 – Параметры модели WSN

Параметр	Значение
Количество узлов	100
Количество кластеров	5
Используемая площадь	(100×100) $м^2$
Размер пакета данных	500 байт
Размер заголовка пакета	25 байт
Протокол маршрутизации	LEACH
Тип трафика	constant bitrate (CBR)
MAC-протокол	CDMA/TDMA
Время симуляции	1 час
Начальная энергия (Дж)	5 , 50
Интенсивность атак	10%, 30%, 50%

От авторов WSN-DS был получен данный датасет, содержащий 19 параметров вместо 23 первоначальных, но, тем не менее, их оказалось достаточно. WSN-DS основан на протоколе LEACH, рассмотрим его подробнее. Узлы сенсорной сети самостоятельно организуются в кластеры и выбирают главу кластера (Cluster Head, CH). На начальном этапе каждый узел с определенной вероятностью, более подробно описанной в [34], и на основе заданной плотности CH в сети предлагает себя в качестве CH.

Каждый CH широковещательно посылает сообщение об объявлении себя CH (ADV-сообщение). На основе силы сигнала каждый узел далее выбирает, к какому CH подключиться, затем отправляет выбранному CH запрос на присоединение [73]. После объединения узлов в кластеры каждый CH создает

TDMA-расписание, позволяющее предотвращать коллизии. Узлы, не являющиеся CH, в свое отведённое время передают данные CH. CH принимает данные, обрабатывает их и передаёт на базовую станцию.

Периодически осуществляется перекластеризация узлов и смена CH для перераспределения энергетической нагрузки. Таким образом, один цикл настройки взаимодействия и непосредственно взаимодействия составляет один раунд. Параметры, представленные в полученном WSN-DS, приведены в Таблице 4.12. Девятнадцатый параметр представляет собой метку о типе атаки характеризуемой другими восемнадцатью параметрами.

Таблица 4.12 – Параметры WSN-DS [106]

Номер параметра	Обозначение параметра	Описание параметра
1	ID	идентификатор узла с указанием номера этапа и раунда
2	Time	текущее время моделирования узла
3	Is_CH	является ли узел CH
4	Who_CH	идентификатор CH
5	Dist_ To_CH	расстояние между узлом и главой кластера
6	ADV_S	количество ADV-сообщений, отправленных узлам
7	ADV_R	количество ADV-сообщений, полученных от CH
8	JOIN_S	количество запросов на присоединение, отправленных узлами главе кластера
9	JOIN_R	количество запросов на присоединение, полученных главой кластера от узлов
10	SCH_S	количество широковещательных координационных сообщений TDMA, отправленных узлам
11	SCH_R	количество широковещательных координационных сообщений TDMA, полученных от CH
12	Rank	порядок узла в расписании TDMA
13	DATA_S	количество пакетов данных, отправленных текущим сенсором главе кластера
14	DATA_R	количество пакетов данных, полученных от CH
15	Data_Sent_ To_BS	количество пакетов данных, отправленных базовой станции
16	dist_CH_ To_BS	расстояние между главой кластера и базовой станцией
17	send_code	код отправки
18	Consumed Energy	количество потребленной энергии
19	Attack type	тип атаки

Датасет WSN-DS содержит 4 вида атак:

– атака «черная дыра» (Blackhole) – предполагает объявление атакующим себя главой кластера, после чего любой узел, подключенный к данному каналу, будет передавать данные атакующему для их отправки на базовую станцию (Base Station, BS), атакующий получает данные, но не отправляет их на BS;

– атака «серая дыра» (Grayhole) – похожа на атаку «черная дыра», но в данном случае часть данных всё же передаётся на BS;

– атака планирования, расписания (Scheduling, атака на TDMA) – предполагает реализацию некорректного конфигурирования работы по протоколу LEACH с целью возникновения коллизий;

– атака «переполнение» (Flooding) – заключается в отправке множества различных сообщений об объявлении CH, в том числе узлам, находящимся далеко, для повышенного расхода ими ресурсов [106].

Кроме сетевых параметров, авторами также оценивалось потребление энергии, полученные результаты представлены в Таблице 4.13 [107].

Таблица 4.13 – Потребление энергии на каждом раунде LEACH во время атаки Flooding

Номер раунда	Интенсивность атаки		
	10%	30%	50%
1	22,03	17,81	18,50
2	25,59	14,56	131,15
3	10,56	128,52	139,06
4	119,50	129,26	131,04
5	17,30	145,68	82,42
6	27,85	39,10	7,54
7	185,07	11,32	78,48
8	156,89	82,78	18,15
9	38,70	15,54	75,79
10	3,79	90,81	–
11	20,77	67,54	–
12	57,09	–	–

Схема обучения двухуровневой ИИС с использованием датасета WSN-DS представлена на Рисунке 4.8. Были проведены вычислительные эксперименты при реализации двухуровневой ИИС в виде распределенной системы с двумя сетями, двумя граничными агентами верхнего уровня и четырьмя агентами нижнего уровня (по два в каждой сети). Данные WSN-DS были нормализованы способом, рассмотренным в разделе 2.2, но так, чтобы значение каждого его параметра не превышало 64 (вместо 255), это позволит значительно уменьшить диапазон значений детекторов и данных, и снизит вычислительную нагрузку, конечно, в случае получения приемлемых результатов анализа.

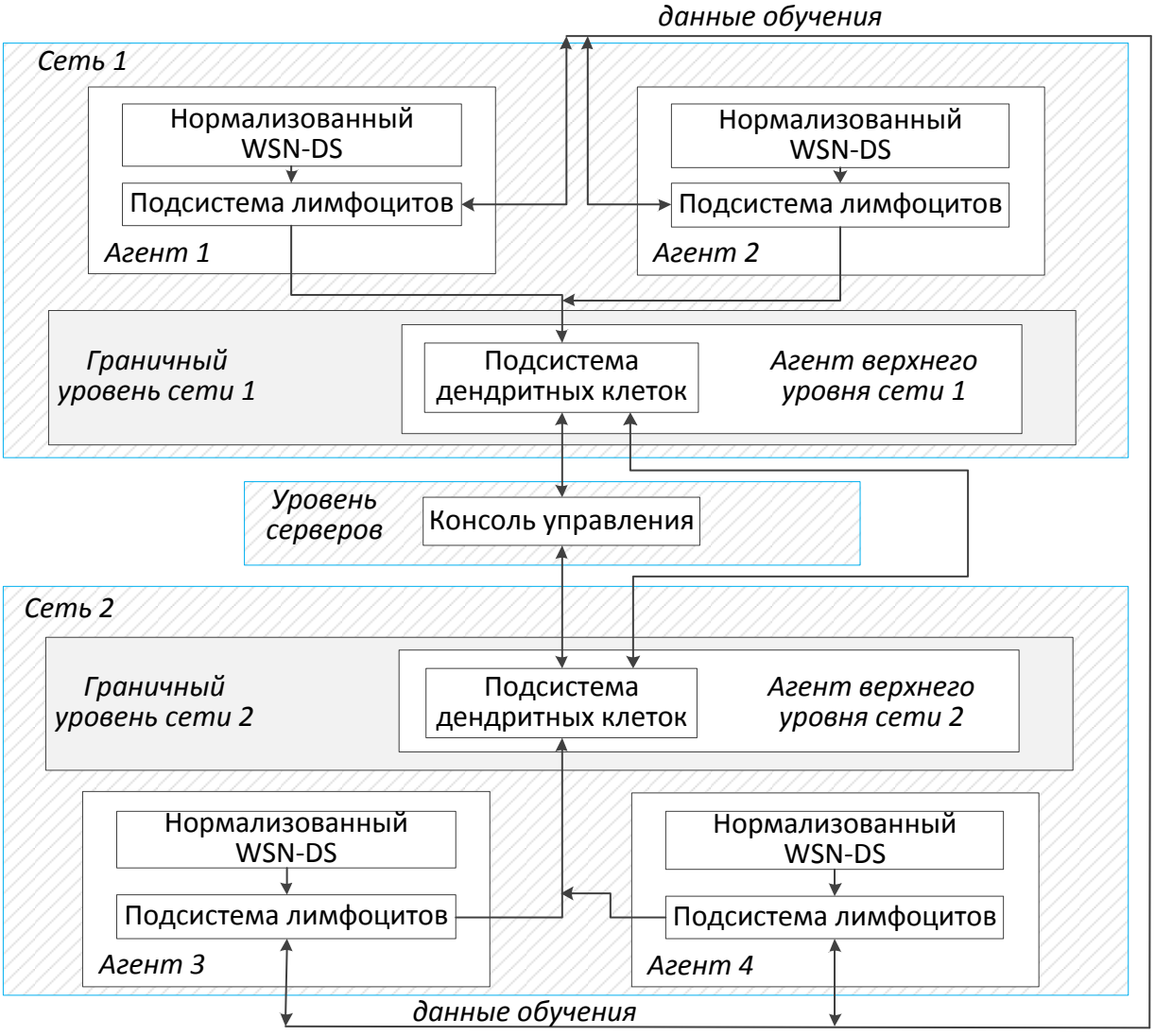


Рисунок 4.8 – Общая схема обучения двухуровневой ИИС с использованием датасета WSN-DS

Параметры были ранжированы по наименьшему проценту совпадений между множествами нормальной и аномальной активности. Результаты сведены в Таблице 4.14. В первую очередь, необходимо было определить рациональное количество параметров для использования в проектируемой системе. Для этого проводился ряд вычислительных экспериментов по аналогии с анализом NSL-KDD с разным количеством параметров, но в режиме, при котором все атаки для системы являются неизвестными.

Таблица 4.14 – Ранжированные параметры WSN-DS

Номер параметра	Обозначение параметра	Процент совпадений по параметру
2	Time	20,65%
18	Consumed Energy	25,69%
15	Data_Sent_To_BS	46,04%
7	ADV_R	50,11%
16	dist_CH_To_BS	60,76%
1	ID	61,15%
9	JOIN_R	78,06%
4	Who_CH	81,82%
10	SCH_S	83,14%
14	DATA_R	86,59%
6	ADV_S	90,46%
11	SCH_R	98,96%
12	Rank	99,18%
5	Dist_ To_CH	99,34%
13	DATA_S	99,82%
17	send_code	99,94%
3	Is_CH	100,00%
8	JOIN_S	100,00%

Уровень ошибок первого рода при каждой эпохе анализа, благодаря алгоритму негативной селекции, не превышал 0,001. Уровень ошибок второго рода, в зависимости от эпохи анализа и количества анализируемых параметров, представлен в Таблице 4.15, по данным которой при применении менее 14 параметров, по аналогии с NSL-KDD, наблюдается предел снижения ошибок второго рода, выделенный серым цветом. Наиболее рациональным оказывается использование 14 параметров. Полученные результаты опубликованы в [190].

Таблица 4.15 – Значения уровня FNR, полученные при анализе WSN-DS

Номер эпохи	Количество параметров						
	8	9	10	11	12	13	14
1	0,999	0,998	0,999	0,999	0,999	0,986	0,992
10	0,664	0,557	0,592	0,382	0,456	0,665	0,523
20	0,459	0,400	0,431	0,299	0,352	0,559	0,280
30	0,317	0,290	0,317	0,255	0,190	0,415	0,221
40	0,215	0,230	0,248	0,198	0,166	0,250	0,172
50	0,151	0,170	0,152	0,140	0,098	0,213	0,111
60	0,088	0,072	0,045	0,099	0,075	0,131	0,079
70	0,050	0,041	0,039	0,047	0,052	0,073	0,069
80	0,050	0,041	0,039	0,018	0,016	0,015	0,047
90	0,050	0,041	0,039	0,018	0,016	0,015	0,020
100	0,050	0,041	0,039	0,018	0,016	0,015	0,008

Здесь и в предыдущих вычислительных экспериментах в качестве метрики расстояния между векторами использовалось расстояние Хэмминга, применяемое в ИИС чаще всего. Для сравнения был проведён ряд дополнительных экспериментов с использованием двух других метрик – косинусной меры и Евклидова расстояния. Результаты, полученные с использованием Евклидова расстояния, оказались в целом аналогичны и сопоставимы с результатами, полученными на основе расстояния Хэмминга. Однако, как показывает практика, время вычисления расстояния Хэмминга более чем в 5 раз меньше времени вычисления Евклидова расстояния.

При использовании косинусной меры близости, эффективность ИИС оказалась крайне низкой. Поэтому был проведен следующий эксперимент. Нормализованный набор данных был разделен на подмножество данных о нормальном состоянии N и подмножество данных об атаках A. Затем для каждого вектора данных N_i с использованием косинусной меры был найден максимально похожий на него вектор A_j. Оказалось, что более чем для 65% таких пар векторов значение косинусной меры превышает 0,98, для более чем 80% – превышает 0,95, хотя число идентичных строк между двумя подмножествами менее 0,01%.

Таким образом, подмножества трудно разделимы косинусной мерой, а применение расстояния Хэмминга оказывается более рациональным, чем

Евклидовой меры или косинусного расстояния. Данные результаты опубликованы в [32].

Далее был проведен ряд вычислительных экспериментов с обучением двухуровневой ИИС на основе 14 параметров нормализованного датасета WSN-DS с применением расстояния Хэмминга и построением ее в виде распределенной системы в соответствии с Рисунком 2.11. Анализ выполнялся агентами нижнего уровня, поочерёдно в целях наблюдения за самообучением агентов. В первой серии экспериментов все атаки были незнакомы системе. Эксперименты показали, что первая эпоха анализа, выполненная первым хостом, выявила всего несколько сотен атак из более чем 30 тысяч, как представлено на Рисунке 4.9

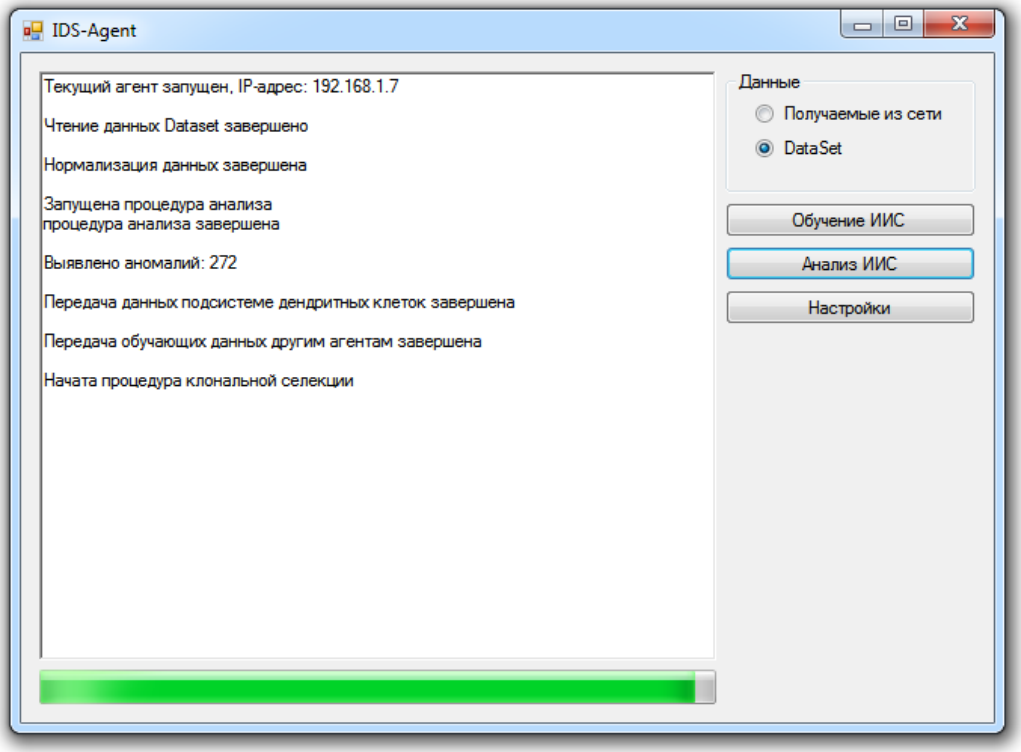


Рисунок 4.9 – Работа первого агента нижнего уровня

Далее среагировавшие детекторы (лимфоциты) были отправлены другим агентам для клональной селекции. Вторая эпоха анализа выполнялась вторым агентом. Это была вторая эпоха для системы в целом, но первая для агента №2. Было выявлено уже больше атак – 1973, как представлено Рисунком 4.10.

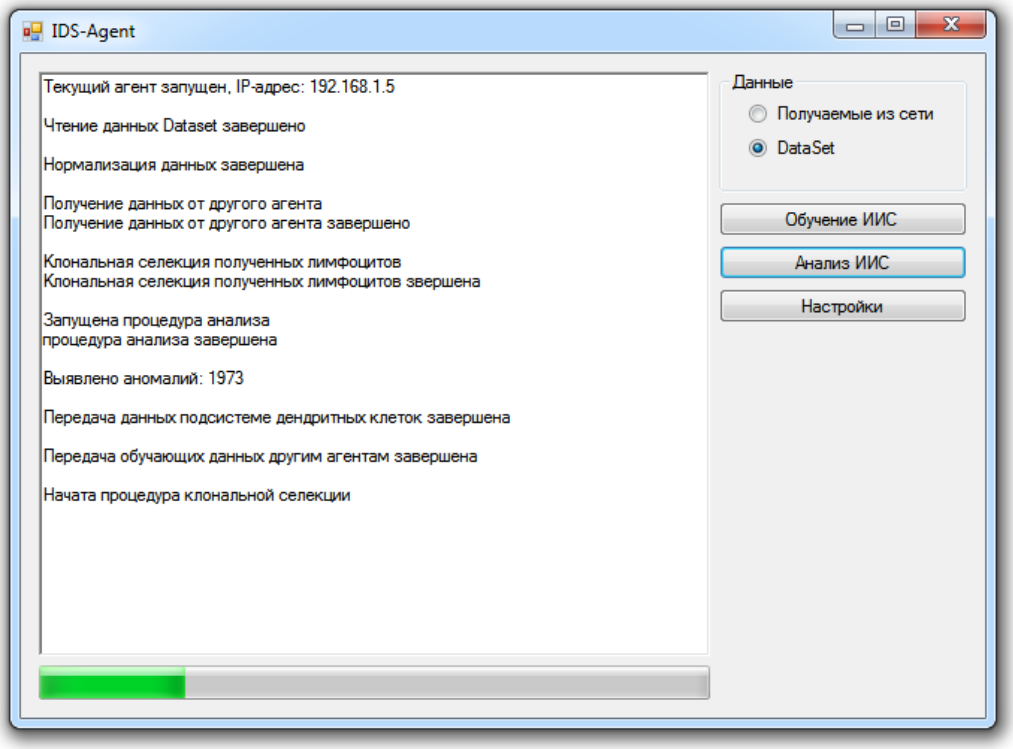


Рисунок 4.10 – Работа второго агента нижнего уровня

Работа третьего и четвертого агентов представлена Рисунками 4.11-4.12.

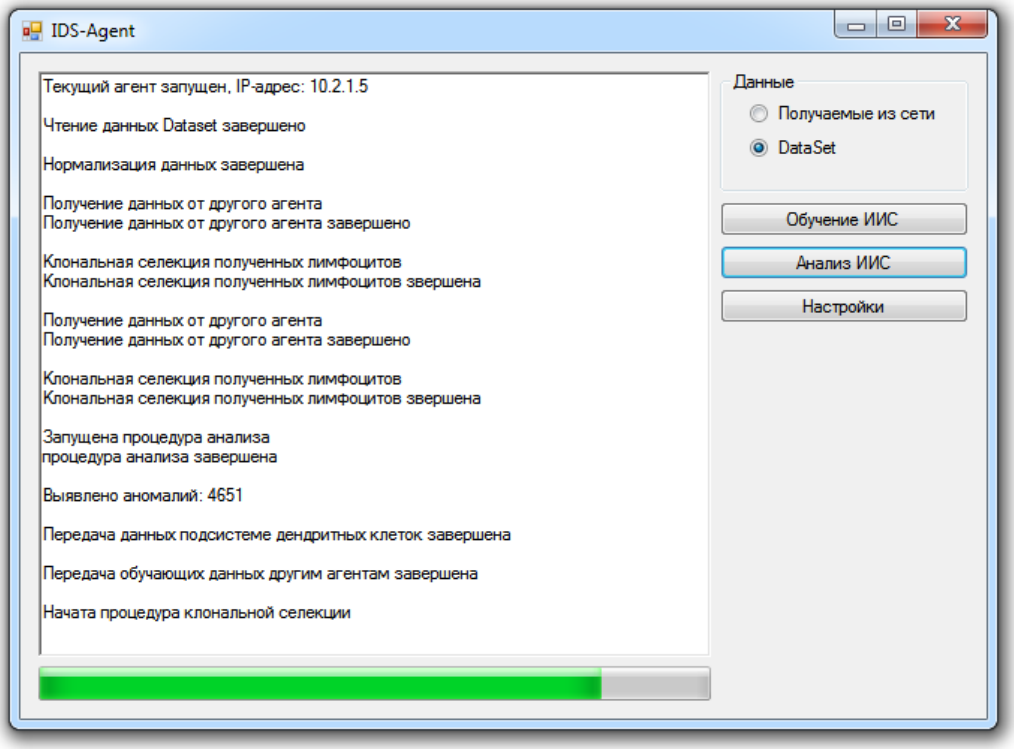


Рисунок 4.11 – Работа третьего агента нижнего уровня

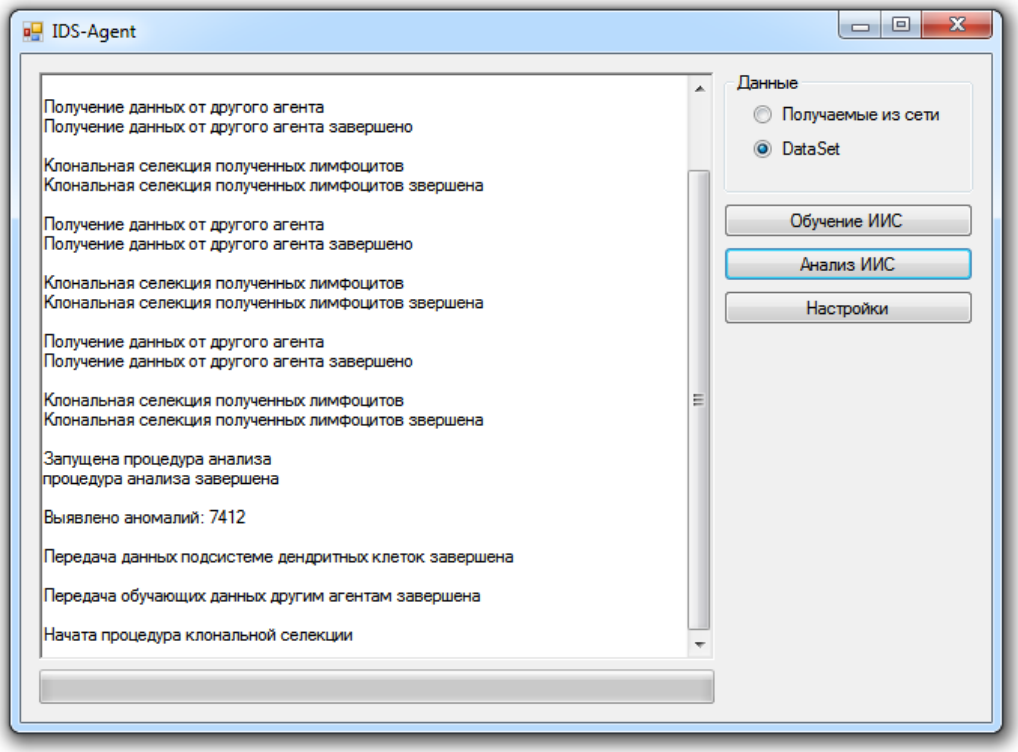


Рисунок 4.12 – Работа четвёртого агента нижнего уровня

Четвертый агент, впервые выполняя анализ, обнаружил уже 7412 атаки против 272, обнаруженных первым агентом, благодаря обучающим данным, полученным от первых трёх эпох анализа, выполненных агентами 1-3. Уровни ошибок первого и второго рода представлены на Рисунке 4.13.

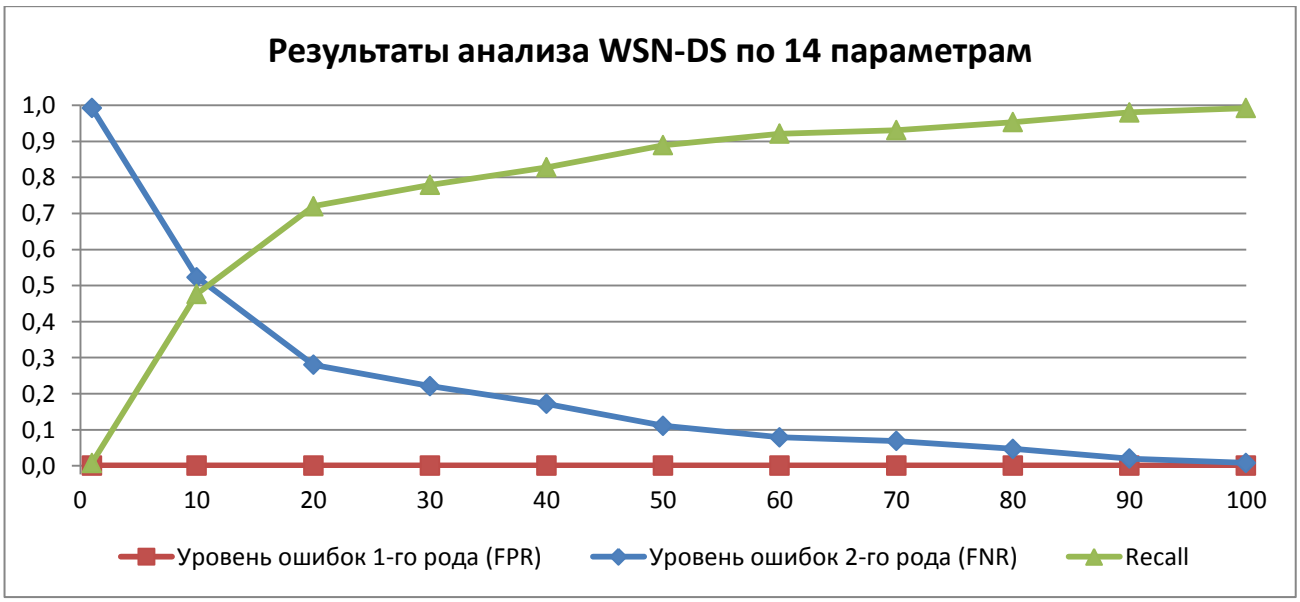


Рисунок 4.13 – Результаты анализа WSN-DS. Ошибки первого и второго рода

Была проведена вторая серия экспериментов следующим образом. ИИС для дополнительного анализа данных WSN-DS, по аналогии с анализом NSL-KDD, была переобучена с применением половины данных об атаках, с целью возможности их классификации. Количество обнаруженных атак для каждого класса атаки относительного общего количества атак данного класса (Recall), поданного на вход агентов нижнего уровня распределенной двухуровневой ИИС представлены в Таблице 4.16.

Таблица 4.16 – Доля обнаруженных атак агентами нижнего уровня

Показатель	Blackhole	Flooding	Grayhole	TDMA	Все атаки, без разделения на классы
Recall	1,00	0,97	0,99	0,97	0,99

Достигнутые значения показателей эффективности для агентов ИИС нижнего уровня в обеих сериях экспериментов представлены в Таблице 4.17.

Таким образом, множество агентов распределенных по сетям эффективно функционируют, совместно обучая друг друга. Также они способны функционировать независимо: если какой-либо из агентов выходит из строя, все остальные продолжают свое функционирование.

Таблица 4.17 – Показатели эффективности обучения ИИС на датасете WSN-DS

№ серии экспериментов	FNR	FPR	TNR	TPR (Recall)	Precision	Accuracy	F_1 score
1	0,005	<0,001	0,999	0,995	0,997	0,999	0,996
2	0,002	<0,001	0,999	0,998	0,999	0,999	0,998

Полученные результаты показывают высокую эффективность выявления атак, в том числе неизвестных, при использовании 14 нормализованных параметров датасета WSN-DS и расстояния Хэмминга, при значении аффинности, равном 12.

В рамках данных экспериментов также оценивалась работа агентов верхнего уровня. В первой серии экспериментов был установлен порог

определения опасности, равный 110 единицам, достигаемый при обнаружении более 10 однотипных аномалий. Первый агент верхнего уровня был подключен к сети 1, анализируя данные двух агентов нижнего уровня, с которых за первые две эпохи поступило 2245 образцов аномалий. Консоль администрирования выводила данные, представленные в Приложении В (Рисунок В.1). По десять первых образцов аномалий каждой группы были отброшены, оставшиеся 2205 были определены как возможно неизвестные атаки и уровень опасности увеличился, он превысил 160 тысяч единиц. Случайные аномалии были бы проигнорированы, но их высокая частота возникновения и схожесть позволяют делать вывод о наличии атаки.

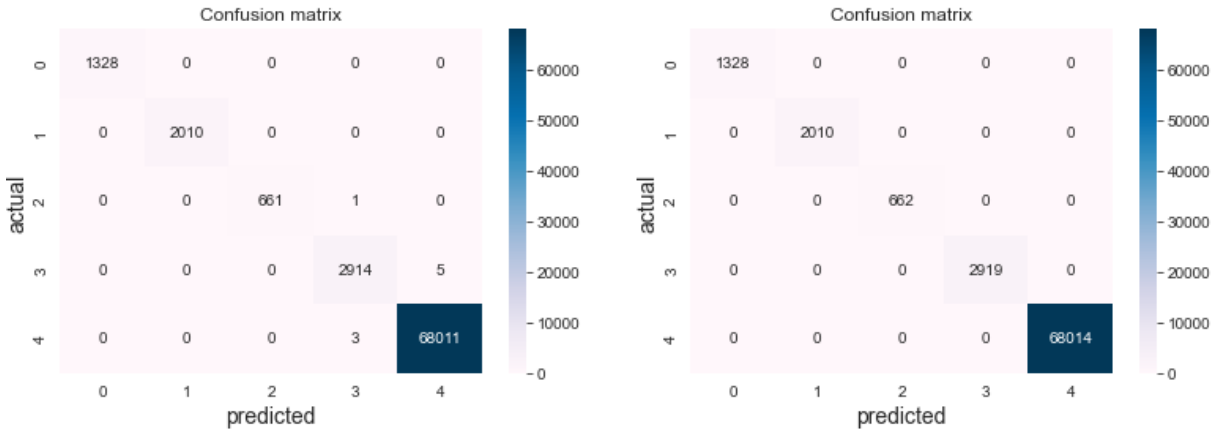
Во второй серии экспериментов на основе WSN-DS пороговое значение уровня опасности было уменьшено до 3 во избежание больших чисел, так как каждая известная атака увеличивает уровень опасности на значение порога. Данные подавались уже обученным агентам нижнего уровня не в полном объеме, а двумя малыми порциями: меньшей – в сети 1, затем большей – в сети 2.

В Приложения В на Рисунке В.2 представлено состояние подсистемы ДК 1 в сети 1, на Рисунке В.3 – подсистемы ДК 2 в сети 2. Данные обо всех выявленных атаках отображаются на консоли. Уровень опасности сети 1 соответствует трёхкратной сумме количества атак сети 1. В сети 2 появляется дополнительный показатель: Other danger events – другие опасные события. Здесь предполагается вывод информации о количестве опасных событий в соответствии с задаваемыми правилами анализа данных SIEM и прочих источников.

На текущий момент данный показатель отличен от нуля, так как вторая система ДК оповещена о превышении порога опасности в первой системе ДК, что является опасным событием. Здесь для данного события задан вес 125 ед.

Далее проводились вычислительные эксперименты по оценке эффективности применения РИСМ для обнаружения атак на WSN с использованием датасета WSN-DS.

Были построены классификаторы ИНС и СЛ, проведены вычислительные эксперименты, матрицы ошибок представлены на Рисунке 4.14.



а) искусственная нейронная сеть *б) случайный лес*

Рисунок 4.14 – Матрицы ошибок при использовании ИНС и СЛ для датасета WSN-DS

Значения показателей точности обнаружения атак при построении комитета классификаторов на основе ИИС, ИНС и СЛ представлены в Таблице 4.18.

Таблица 4.18 – Показатели точности обнаружения атак WSN-DS при использовании комитета классификаторов

Мера	FNR	FPR	TNR	TPR (Recall)	Precision	Accuracy	F_1 score
Значение для комитета классификаторов	0,001	0,001	0,999	0,999	0,999	0,999	0,999

Как видно, показатели точности обнаружения атак в составе датасета WSN-DS с помощью комитета классификаторов аналогичны результатам, полученным для датасетов NSL-KDD и WUSTL-IIOT-2021, что дополнительно подтверждает эффективность системы. При обнаружении с помощью ИИС неизвестной атаки, она классифицируется как неизвестная и передается дальше в ИНС и СЛ. Если последние однозначно сходятся во мнениях о классе атаки, ей присваивается один из известных классов, если мнения ИНС и СЛ отличаются, атака так и остаётся в классе неизвестных.

Преимущества рассмотренного подхода заключаются:

– в удобстве интеграции системы в распределенную сеть промышленного Интернета вещей;

– в обеспечении высокого уровня обнаружения известных и неизвестных сетевых атак с настройкой чувствительности реагирования на одиночные аномалии, что позволяет избегать параноидального реагирования на каждую аномалию;

– в обеспечении высокой точности классификации известных атак.

В целом, предлагаемая РИСМ демонстрирует высокие значения показателей эффективности обнаружения сетевых атак и аномалий в системах промышленного Интернета вещей, а также классификации атак.

Выводы по главе 4.

1. Разработаны архитектура многоагентной распределенной интеллектуальной систем мониторинга (РИСМ) ИБ промышленного Интернета вещей включающая три уровня, первые два из которых функционируют на граничном уровне архитектуры IIoT, третий – на уровне платформы архитектуры IIoT; методика обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием РИСМ.

2. Эффективность предлагаемой РИСМ оценивалась на тестовой IIoT-системе контроля уровня и мутности воды в резервуаре, которая, в свою очередь, входит в состав автоматизированной системы очистки и распределения воды в промышленных резервуарах. Проводились вычислительные эксперименты по обнаружению атак на основе датасета WUSTL-IIOT-2021 созданного его авторами на основе сетевого взаимодействия элементов указанной IIoT-системы, в том числе во время имитированных сетевых атак. Рассматривались несколько вариантов построения ИИС, ИНС, СЛ, оценивались значения показателей эффективности отдельных алгоритмов и комитета классификаторов. Совокупность показателей эффективности комитета классификаторов оказалась

выше. По метрикам: Precision, Recall, Accuracy, F_1 score, TNR, TPR достигнуты значения 0,999 на основе тестового набора данных.

3. РИСМ была протестирована на основе данных испытательного стенда беспроводной сенсорной сети, состоящей из 100 сенсорных узлов, объединенных в 5 кластеров. На основе взаимодействия элементов этой сети, в том числе во время имитированных сетевых атак, авторами стенда был создан датасет WSN-DS, содержащий атак на протокол LEACH, часто используемый в WSN. Проведен ряд вычислительных экспериментов по оценке эффективности РИСМ в выявлении атак WSN-DS, в том числе с применением различных мер близости в ИИС. Результаты экспериментов показали рациональность применения расстояния Хэмминга, а также комитета классификаторов: ИИС, ИНС, СЛ, продемонстрировавшего высокую эффективность в обнаружении и классификации атак на WSN.

Заключение

1. В работе проведен анализ современного состояния исследований в области мониторинга сетей промышленного Интернета вещей с использованием технологий искусственного интеллекта, предложена концепция построения интеллектуальной системы мониторинга ИБ промышленного Интернета вещей с использованием механизмов искусственных иммунных систем, методов машинного обучения, взаимодействия с подсистемой корреляции событий информационной безопасности (SIEM-системой), что позволило повысить полноту и точность выявления внешних и внутренних угроз информационной безопасности.

2. Разработаны алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей на основе механизмов искусственных иммунных систем (ИИС), реализованные в составе двухуровневой распределенной ИИС, состоящей из множества взаимодействующих агентов, интегрирующих различные подходы в рамках теории ИИС (негативная селекция, клональный отбор, теория опасности, теория иммунной сети) и модификации известных алгоритмов клональной селекции, дендритных клеток и обновления детекторов, что позволяет в совокупности существенно снизить уровень принятия ошибочных решений, выявлять в том числе новые, ранее неизвестные системе угрозы безопасности информации.

3. Разработаны алгоритмы обнаружения атак и аномалий сетевого трафика промышленного Интернета вещей с использованием ансамбля методов машинного обучения, включающего искусственную иммунную систему, искусственную нейронную сеть, алгоритм случайного леса, с интеграцией подсистемы корреляции событий ИБ, реализующего трёхуровневый интеллектуальный анализ данных сетевого трафика с целью определения наличия сетевых атак и аномалий с учетом уровня опасности, с возможностью корректировки чувствительности, что позволило достичь более высоких значений

показателей эффективности в сравнении с применением данных алгоритмов по отдельности.

4. Разработана архитектура исследовательского прототипа распределенной интеллектуальной системы мониторинга (РИСМ) ИБ промышленного Интернета вещей, состоящей в классе многоуровневых многоагентных гибридных систем, что облегчает её интеграцию в различные подсети IIoT непосредственно, позволяет использовать преимущества не только централизованной обработки данных, но и локальных, а также граничных вычислений, обеспечивает возможность интеграции с существующими систем безопасности и мониторинга, включая межсетевые экраны, SIEM, SCADA и др.

Эффективность предлагаемой РИСМ оценивалась на тестовой IIoT-системе контроля уровня и мутности воды в резервуаре, на основе датасета WUSTL-IIOT-2021, а также на основе данных сетевого взаимодействия беспроводной сенсорной сети, сведенных в датасет WSN-DS. Значения показателей эффективности РИСМ на тестовых наборах данных были на уровне 98-99%, что превышает значения показателей эффективности существующих систем в среднем на 1,5% и обеспечивает высокий уровень обнаружения сетевых атак и аномалий, демонстрирует высокую эффективность системы.

Список сокращений и условных обозначений

АСУ	Автоматизированная система управления
АСУ ТП	Автоматизированная система управления технологическим процессом
ВРТП	Временные ряды технологических параметров
ГА	Генетический алгоритм
ГИС	Гибридная интеллектуальная система
ГИСОА	Гибридная интеллектуальная система обнаружения атак
ДК	Дендритная клетка
ДМ	Датчик мутности
ДР	Дерево решений
ДС	Датасет
ДУ	Датчик уровня
ЕИС	Естественная иммунная система
ИД	Источник данных
ИИ	Искусственный интеллект
ИИС	Искусственная иммунная система
ИНС	Искусственная нейронная сеть
ИТКС	Информационно-телекоммуникационная сеть
ИУС	Информационно-управляющая система
КАД	Корреляционный анализ данных
КИИ	Критическая информационная инфраструктура
КФС	Киберфизическая система
МАС	Многоагентная система
МЭ	Межсетевой экран
НБК	Наивный байесовский классификатор
ОС	Операционная система
ПО	Программное обеспечение
СЛ	Случайный лес
СОА	Система обнаружения атак
ТП	Технологический процесс
ЦОД	Центр обработки данных
BS	Base Station – базовая станция
CH	Cluster head – глава кластера
CSA	Clonal Selection Algorithm – Алгоритм клональной селекции
DoS	Denial of Service – отказ в обслуживании
HMI	Human Machine Interface – человеко-машинный интерфейс
IIoT	Industrial Internet of Things – Промышленный интернет вещей.
IoT	Internet of Things – Интернет вещей
KNN	k-nearest neighbors – алгоритм k ближайших соседей
NK	Natural Killer – естественный киллер
NSA	Negative Selection Algorithm –
PLC	Programmable Logic Controller – программируемый логический контроллер
SCADA	Supervisory Control And Data Acquisition – система диспетчерского управления и сбора данных
SIEM	Security Information and Event Management – система управления безопасностью и событиями безопасности
SVM	Support Vector Machine – машина опорных векторов
WSN	Wireless Sensor Network –беспроводная сенсорная сеть

Словарь терминов

Аномалия (сетевого трафика) – существенное отклонение трафика сетевого устройства от нормального профиля трафика для данного устройства или группы устройств АСУ ТП.

Апплет – несамостоятельный компонент программного обеспечения, работающий в контексте другого, полновесного приложения.

Атака – умышленное посягательство на систему, продуманная попытка обойти сервисы безопасности и нарушить политику безопасности системы, попытка реализации угрозы информационной безопасности посредством программных или программно-аппаратных средств.

Аффинность – мера схожести между анализируемым экземпляром данных и детектором, используемая в искусственной иммунной системе.

Беспроводная сенсорная сеть (Wireless Sensor Network, WSN) – сеть, состоящая из большого числа автономных сенсорных узлов, собирающих различные данные и обменивающихся ими при помощи беспроводного соединения с более мощным узлом – базовой станцией.

Датасет – собой совокупность данных о нормальных сетевых взаимодействиях и соединениях, характерных атакам, представленных построчно с указанием класса, к которому относится конкретная строка, предназначенная для обучения и тестирования систем обнаружения атак.

Дендритная клетка – модель, используемая в искусственной иммунной системе, отвечающая за анализ опасности и снижение реагирования на выявленные неопасные аномалии.

Интернет вещей (Internet of Things, IoT) – инфраструктура взаимосвязанных систем, сущностей, служб и информационных ресурсов, служащих для обработки информации о физическом и виртуальном мире и реагирования на нее. Под сущностью здесь понимается обособленно существующий предмет

Информационно-управляющая система (Автоматизированная система управления) – комплекс программных и аппаратных средств, предназначенных для реализации функций управления.

Киберфизическая система – интеллектуальная система, включающая в себя взаимодействующие инженерные сети физических и вычислительных компонентов.

Клональная селекция – один из основных алгоритмов искусственной иммунной системы, обеспечивающий адаптивность системы, самообучение лучшему обнаружению аномалий и атак, подобных выявленным ранее.

Критическая информационная инфраструктура – объекты критической информационной инфраструктуры и сети электросвязи, которые используются для организации взаимодействия таких объектов.

Лимфоцит – 1) клетка естественной иммунной системы, осуществляющая распознавание и уничтожение чужеродного патогена; 2) модель, используемая в искусственной иммунной системе для обнаружения среди анализируемых образцов элементов, соответствующих аномальному состоянию, контролируемой системы.

Негативная селекция (отрицательный отбор) – один из основных алгоритмов искусственной иммунной системы, обеспечивающий снижение частоты ложных реакций детекторов на нормальное состояние контролируемой системы.

Объекты критической информационной инфраструктуры – информационные системы, автоматизированные системы управления, информационно-телекоммуникационные сети субъектов, функционирующих в сфере здравоохранения, связи, науки, энергетики, транспорта, в области атомной энергии, ракетно-космической, оборонной, металлургической горнодобывающей и химической промышленности, банковской сфере и иных сферах финансового рынка, топливно-энергетического комплекса.

Промышленная сеть управления – сеть, предназначенная для обеспечения связи оборудования промышленной системы управления для

реализации вычислений и управления производственными процессами, интегрированных между собой для управления промышленным производством, распределением ресурсов и передачей информации.

Промышленный интернет вещей (Industrial Internet of Things, IoT) – составная часть Интернета вещей, нацеленная на выполнение промышленных задач, представляет собой многоуровневую информационно-управляющую систему, включающую в себя средства автоматизации технологического процесса, широкий спектр контроллеров, датчиков и исполнительных механизмов, используемых на различных узлах и агрегатах промышленных объектов, средства обработки, передачи и визуализации собираемых данных о состоянии этих объектов, аналитические инструменты интерпретации получаемой информации, поддержки принятия управленческих решений и многие другие компоненты.

Система обнаружения атак – система, анализирующая сетевой трафик, выявляющая образцы, соответствующие атакам (сигнатуры) или несоответствующие нормальному сетевому взаимодействию (аномалии).

Умное производство – область интегрированных ресурсов (в том числе человеческих) для создания и доставки продуктов и услуг, взаимодействующую с другими звеньями цепочки создания стоимости предприятия и улучшает показатели производительности.

Hardening – усиление защищенности системы.

SIEM (Security Information and Event Management) – система управления безопасностью и событиями безопасности, система анализирующая поток данных о событиях безопасности, поступающих с различных устройств в виде записей системных журналов и т.п., агрегирующая данные и выполняющая корреляционный анализ с целью выявления скрытых взаимосвязей между событиями и идентификации инцидентов информационной безопасности.

Список литературы

1. Федеральный закон от 26.07.2017 N 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации» [Электронный ресурс] // КонсультантПлюс. – URL: https://www.consultant.ru/document/cons_doc_LAW_220885/ (дата обращения 01.02.2023).

2. ГОСТ Р 54411-2018/ISO/IEC TR 24722:2015 Информационные технологии. Биометрия. Мультимодальные и другие мультибиометрические технологии [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Data/704/70452.pdf (дата обращения: 24.12.2022).

3. ГОСТ Р 56205-2014 IEC/TS 62443-1-1:2009 Сети коммуникационные промышленные. Защищенность (кибербезопасность) сети и системы. Часть 1-1. Терминология, концептуальные положения и модели [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Index/582/58241.htm (дата обращения: 02.03.2021).

4. ГОСТ Р 59547-2021 Защита информации. Мониторинг информационной безопасности. Общие положения [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Data/762/76267.pdf (дата обращения 05.06.2022).

5. ГОСТ Р ИСО/МЭК ТО 18044-2007 Информационная технология. Методы и средства обеспечения безопасности. Менеджмент инцидентов информационной безопасности [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Data2/1/4293832/4293832382.pdf (дата обращения 14.03.2023).

6. ГОСТ Р МЭК 62443-2-1-2015 IEC 62443-2-1:2010 Сети коммуникационные промышленные. Защищенность (кибербезопасность) сети и системы. Часть 2-1. Составление программы обеспечения защищенности (кибербезопасности) системы управления и промышленной автоматики [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Index/60/60330.htm (дата обращения 26.04.2023).

7. ГОСТ Р МЭК 62443-3-3-2016 IEC 62443-3-3:2013 Сети промышленной коммуникации. Безопасность сетей и систем. Требования к системной безопасности и уровни безопасности [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Data/627/62748.pdf (дата обращения: 02.03.2021).

8. ПНСТ 417-2020 Система киберфизическая. Термины и определения: предварительный национальный стандарт РФ [Электронный ресурс] // База ГОСТов. – URL: https://allgosts.ru/35/110/pnst_417-2020 (дата обращения 01.04.2023).

9. ПНСТ 420-2020 Информационные технологии. Интернет вещей промышленный. Типовая архитектура: предварительный национальный стандарт РФ [Электронный ресурс] // АНО МЦК: центр сертификации и стандартизации. – URL: https://files.stroyinf.ru/Data2/1/4293719/4293719799.pdf (дата обращения 26.04.2023).

10. ПНСТ 643-2022 Информационные технологии. Интернет вещей промышленный. Термины и определения: предварительный национальный стандарт РФ [Электронный ресурс] // База ГОСТов. – URL: https://allgosts.ru/35/020/pnst_643-2022 (дата обращения: 23.05.2023).

11. Приказ ФСТЭК от 14 марта 2014 г. N 31 «Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной сред» [Электронный ресурс] // ФСТЭК РФ. – URL: https://fstec.ru/normotvorcheskaya/akty/53-prikazy/868-prikaz-fstek-rossii-ot-14-marta-2014-g-n-31 (дата обращения 22.05.2022).

12. Приказ ФСТЭК России от 21.12.2017 N 235 «Об утверждении Требований к созданию систем безопасности значимых объектов критической информационной инфраструктуры Российской Федерации и обеспечению их функционирования» [Электронный ресурс] // ФСТЭК РФ. – URL: https://fstec.ru/normotvorcheskaya/akty/53-prikazy/1589-prikaz-fstek-rossii-ot-21-dekabrya-2017-g-n-236 (дата обращения 22.05.2022).

13. Приказ ФСТЭК России от 25.12.2017 N 239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации» [Электронный ресурс] // ФСТЭК РФ. – URL: http://www.consultant.ru/document/cons_doc_LA W_294287/ (дата обращения 22.05.2022).

14. Абрамов Е.С., Басан Е.С., Макаревич О.Б. Разработка системы обнаружения атак для кластерной беспроводной сенсорной сети // Информационное противодействие угрозам терроризма. – 2013. – № 20. – С. 134-140.

15. Абрамова Т.В., Аралбаев Т.З., Аралбаева Г.Г., Галимов Р.Р. Метод оперативного прогнозирования и ранжирования рисков информационной безопасности на основе ассоциативного подхода // Вопросы развития современной науки и практики в период становления цифровой экономики. – СПб.: СПбГЛТУ, 2018. – 300 с.

16. Ажмухамедов И.М., Марьенков А.Н. Поиск и оценка аномалий сетевого трафика на основе циклического анализа // Инженерный вестник Дона. – 2012. – №2. – С. 17-26.

17. Алабугин С.К., Соколов А.Н. Обнаружение вторжений в автоматизированных системах управления технологическими процессами с использованием ансамбля моделей рекуррентной и двунаправленной

генеративно-состязательной нейронных сетей // Вестник УрФО. – 2021. – №3 (41). – С. 38-48.

18. Алейнов Ю.В. Метод повышения эффективности обнаружения сетевых атак неизвестного типа путем внедрения ложных целей в состав сети [Электронный ресурс] // Доклады ТУСУР. – 2014. – №2 (32). – С. 40-43. – URL: https://cyberleninka.ru/article/n/metod-povysheniya-effektivnosti-obnaruzheniya-setevyh-atak-neizvestnogo-tipa-putem-vnedreniya-lozhnyh-tseley-v-sostav-seti (дата обращения: 22.05.2023).

19. Аникин И.В., Емалетдинова Л.Ю., Кирпичников А.П. Методы оценки и управления рисками информационной безопасности в корпоративных информационных сетях [Электронный ресурс] // Вестник технологического университета. – 2015. – Т. 18. – №6. – URL: https://cyberleninka.ru/article/n/metody-otsenki-i-upravleniya-riskami-informatsionnoy-bezopasnosti-v-korporativnyh-informatsionnyh-setyah (дата обращения: 24.05.2023).

20. Атарская Е.А. Система обнаружения аномалий технологических временных рядов параметров промышленного проекта // Мавлютовские чтения: материалы XV Всероссийской молодежной научной конференции: в 7 томах. – Уфа, 2021. – Том 4. – 317-325.

21. Бахарева Н.Ф., Тарасов В.Н., Шухман А.Е., Полежаев П.Н., Ушаков Ю.А., Матвеев А.А. Выявление атак в корпоративных сетях с помощью методов машинного обучения // Современные информационные технологии и ИТ-образование. – 2018. – №3. – С. 626-632.

22. Березин Д. Как сегодня строится центр оперативного управления информационной безопасностью (SOC-центр) [Электронный ресурс] // Хабр. – URL: https://habr.com/ru/company/croc/blog/353324/ (дата обращения 31.05.2020).

23. Борисов М.А., Заводцев И.В. Инструментальные средства оценки уязвимостей в автоматизированных системах // Вестник РГГУ. Серия документоведение и архивоведение. Информатика. Защита информации и информационная безопасность. – 2010. – № 12 (55). – С. 259-262 .

24. Боршевников А.Е. Сетевые атаки. Виды. Способы борьбы // Современные тенденции технических наук: материалы Междунар. науч. конф. (г. Уфа, октябрь 2011 г.). – Уфа: Лето, 2011. – С. 8-13.

25. Брюхомицкий Ю.А. Искусственные иммунные системы в информационной безопасности: уч. пособие / Ю.А. Брюхомицкий; Южный федеральный университет. – Ростов-на-Дону – Таганрог: Издательство Южного федерального университета, 2019. – 142 с.

26. Бурлаков М.Е. Оптимизация атрибутного пространства у L4, L7 наборов данных // Научные труды КубГТУ. – 2022. – №5. – С. 113-125.

27. Бурлаков М.Е., Ивкин А.Н. Система обнаружения вторжения на основе искусственной иммунной системы // Вестник ПНИПУ. – 2019. – № 29. – С. 209-224.

28. Васильев В.И., Вульфин А.М., Гвоздев В.Е., Шамсутдинов Р.Р. Гибридная интеллектуальная система обнаружения атак на основе комбинации методов машинного обучения [Электронный ресурс] // Моделирование,

оптимизация и информационные технологии. – 2021. – №9 (3). – URL: https://moitvivt.ru/ru/journal/article?id=1032 (дата обращения 14.03.2023).

29. Васильев В.И., Вульфин А.М., Гвоздев В.Е., Шамсутдинов Р.Р. Гибридная система обнаружения атак на основе комитета классификаторов [Электронный ресурс] // Моделирование, оптимизация и информационные технологии. – 2022. – Т.10. – №4.– URL: https://moitvivt.ru/ru/journal/pdf?id=1267 (дата обращения 14.03.2023).

30. Васильев В.И., Вульфин А.М., Гвоздев В.Е., Шамсутдинов Р.Р. Комплексирование механизмов искусственных иммунных систем в составе интегрированной системы обнаружения атак на промышленный Интернет вещей // Моделирование, оптимизация и информационные технологии. – 2022. – №10 (4). – URL: https://moitvivt.ru/ru/journal/pdf?id=1240 (дата обращения: 29.04.2023).

31. Васильев В.И., Вульфин А.М., Картак В.М., Кириллова А.Д., Миронов К.В. Система обнаружения атак в беспроводных сенсорных сетях промышленного Интернета // Труды ИСА РАН. – 2019. – Том 69. – №4.– С. 70-78.

32. Васильев В.И., Гвоздев В.Е., Шамсутдинов Р.Р. Обнаружение аномалий в системах промышленного Интернета вещей на основе искусственной иммунной системы // Доклады ТУСУР. – 2021. – №4 (24). – С. 40-45.

33. Васильев В.И., Шамсутдинов Р.Р. Интеллектуальная система анализа инцидентов информационной безопасности (на основе методологии SIEM-систем с применением механизмов иммунокомпьютинга) // Моделирование, оптимизация и информационные технологии. – 2019. – №1 (7). – С. 536-547.

34. Виноградов Г.П., Емцев А.С., Федотов И.С. Беспроводные сенсорные сети в защищаемых зонах // Известия ЮФУ. Технические науки. – 2021. – №1 (218). – С. 19-30.

35. Вульфин А.М. Модели и методы комплексной оценки рисков безопасности объектов критической информационной инфраструктуры на основе интеллектуального анализа данных: диссертация на соискание учёной степени доктора технических наук, Федеральное государственное бюджетное образовательное учреждение высшего образования «Уфимский государственный авиационный технический университет» (УГАТУ). – Уфа, 2022 [Электронный ресурс] // УГАТУ. – URL: https://ugatu.su/media/uploads/MainSite/Science/dissovet/07/2022/vulfin-am/dissert.pdf (дата обращения 29.01.2023).

36. Гайфулина Д.А., Котенко И.В. Анализ моделей глубокого обучения для задач обнаружения сетевых аномалий интернета вещей // Информационно управляющие системы. – 2021. – №1. – С. 28-37.

37. Горюнов М.Н., Мацкевич А.Г., Рыболовлев Д.А. Синтез модели машинного обучения для обнаружения компьютерных атак на основе набора данных CICIDS2017 // Труды ИСП РАН. – 2020. – Т. 32. – №. 5. – С. 81-94.

38. Гузаиров М.Б., Машкина И.В., Тухватшин Т.Х. Разработка моделей принятия решений по оперативному управлению защитой информации на основе

численной оценки вероятности атаки // Известия ЮФУ. Технические науки. – 2008. – № 8(85). – С. 18-24.

39. Дрозд А. Обзор SIEM-систем на мировом и российском рынке [Электронный ресурс] // Anti-malware. – URL: https://www.anti-malware.ru/analytics/Technology_Analysis/Overview_SECURITY_systems_global_and_Russian_market (Дата обращения: 07.12.2022).

40. Дубровин М.Г., Глухих И.Н. Модели и методы проактивного мониторинга ИТ-систем // Моделирование, оптимизация и информационные технологии. – 2018. – №1(20). – С. 314-324.

41. Еременко Ю.И., Глущенко А.И. О решении неформализуемых и плохоформализуемых задач методами иммунных алгоритмов // Информационные технологии. – 2011. – №7. – С. 2-7.

42. Жуков В.Г., Бухтояров В.В. Разработка и исследование метода обнаружения инцидентов информационной безопасности на основе коллективов интеллектуальных информационных технологий // Решетневские чтения. – 2013. – Т. 2. – С. 283-284.

43. Зинкевич А.В., Еремин К.Ю. Система обнаружения вторжений с использованием нейронной сети для анализа данных // Ученые заметки ТОГУ. – 2017. – Т. 8. – №4. – С. 514-519.

44. Зуев В.Н. Обнаружение аномалий сетевого трафика методом глубокого обучения // Программные продукты и системы. – 2021. – №1. – С. 91-97.

45. Информационная безопасность в IoT [Электронный ресурс] // Хабр. – URL: https://habr.com/ru/post/700800/ (дата обращения 06.04.2023).

46. Информационная безопасность цифрового пространства / под ред. Е.В. Стельмашонок, И.Н. Васильевой. – СПб. : Изд-во СПбГЭУ, 2019. – 155 с.

47. Как за один день разработать SIEM (систему управления инцидентами информационной безопасности) [Электронный ресурс] // Хабр. – URL: https://habr.com/ru/articles/353542/ (дата обращения: 13.04.2023).

48. Катасёв А.С., Катасёва Д.В., Кирпичников А.П. Нейросетевая диагностика аномальной сетевой активности [Электронный ресурс] // Вестник технологического университета. – 2015. – Т.18. – №6. – С. 163-167. – URL: https://cyberleninka.ru/article/n/neyrosetevaya-diagnostika-anomalnoy-setevoy-aktivnosti (дата обращения 12.05.2023).

49. Кибербезопасность промышленных предприятий под угрозой [Электронный ресурс] // ИКС Медиа. – URL: http://www.iksmedia.ru/news/5394915-Kiberbezopasnost-promyshlennyx-pred.html (дата обращения: 12.03.2018).

50. Костогрызов А.И., Лазарев В.М., Любимов А.Е. Прогнозирование рисков для обеспечения эффективности систем информационной безопасности в их жизненном цикле // Правовая информатика. – 2013. – №4. – С4-16.

51. Котенко И.В., Саенко И.Б., Дойникова Е.В. Оценка рисков в компьютерных сетях критических инфраструктур // Инновации в науке. – 2013. – № 16-1. – С. 84-88.

52. Котенко И.В., Саенко И.Б., Лаута О.С., Крибель А.М. Методика обнаружения аномалий и кибератак на основе интеграции методов фрактального анализа и машинного обучения // Информатика и автоматизация. – 2022. – Т. 6. – №. 21. – С. 1328-1358.

53. Котов В.Д., Васильев В.И. Система обнаружения сетевых вторжений на основе механизмов иммунной модели // Известия ЮФУ. Технические науки. – 2011. – № 12 (125). – С. 180-189.

54. Кушнир Е. Протоколы интернета вещей: как обмениваются данными IoT-устройства, серверы и пользовательские приложения [Электронный ресурс] // VK Cloud : журнал об IT-бизнесе, технологиях и цифровой трансформации. – URL: https://mcs.mail.ru/blog/protokoly-interneta-veschej (дата обращения: 11.06.2023).

55. Лаборатория Касперского, Что угрожает промышленному интернету вещей и как от этого защититься [Электронный ресурс] // Vc.ru. – URL: https://vc.ru/kaspersky/265770-chto-ugrozhaet-promyshlennomu-internetu-veshchey-i-kak-ot-etogo-zashchititsya (дата обращения: 30.07.2021).

56. Лаборатория Касперского: распространение умных устройств в промышленности повлечёт за собой смену подхода к киберзащите [Электронный ресурс] // Лаборатория Касперского. – URL: https://www.kaspersky.ru/about/press-releases/2020_laboratoriya-kasperskogo-rasprostranenie-umnih-ustroistv-v-promishlennosti-povlechyot-za-soboi-smenu-podhoda-k-kiberzaschite. (дата обращения: 30.07.2021).

57. Лаврова Д.С., Полтавцева М.А., Печенкин А.И., Зегжда Д.П. SIEM-система для обнаружения и анализа инцидентов безопасности в интернете вещей // Методы и технические средства обеспечения безопасности информации. – 2016. – №25. – С. 35-36.

58. Лившиц И.И. Менеджмент информационной безопасности // Стандарты и качество. – 2017. – №9. – С. 48-52.

59. Магницкий Н.А. Использование иммунной сети для обнаружения атак на ресурсы распределенных информационных систем // Информационные технологии и вычислительные системы. – 2009. – №3. – С. 22-26.

60. Машкина И.В., Сенцова А.Ю. Обеспечение информационной безопасности системы облачных вычислений // Информационные технологии. – 2016. – Т. 22. – № 11.– С. 843-853.

61. Меньшов М. Коэффициент корреляции Пирсона [Электронный ресурс] // Казанский федеральный университет. – URL: https://kpfu.ru/portal/docs/F_2064674290/NPS_19.Pirson.Menshov.pdf (дата обращения: 13.04.2023).

62. Мещеряков Р.В., Ходашинский И.А., Гусакова Е.Н. Оценка информативного признакового пространства для системы обнаружения вторжений // Известия ЮФУ. Технические науки. – 2013. – № 12 (149). – С. 57-63.

63. Микова С.Ю., Оладько В.С., Нестеренко М.А. Подход к классификации аномалий сетевого трафика // Инновационная наука. – 2015. – № 11. – С. 78-81.

64. Милославская Н.Г. Центры управления информационной безопасностью // Безопасность информационных технологий. – 2016. – № 4(23). – С. 38-51.

65. Обнаружение аномалий в данных сетевого мониторинга методами статистики [Электронный ресурс] // Хабр. – URL: https://habr.com/ru/post/344762/ (дата обращения: 12.03.2018).

66. Орешкина Д. Эталонная архитектура безопасности интернета вещей (IoT). Часть 1 [Электронный ресурс] // Anti-malware. – URL: https://www.anti-malware.ru/practice/solutions/iot-the-reference-security-architecture-part-1 (дата обращения 06.03.2023).

67. Орешкина Д. Эталонная архитектура безопасности интернета вещей (IoT). Часть 2 [Электронный ресурс] // Anti-malware. – URL: https://www.anti-malware.ru/practice/solutions/iot-reference-architecture-protection-part-2 (дата обращения 06.03.2023).

68. Основные метрики задач классификации в машинном обучении [Электронный ресурс] // Webiomed : платформа прогнозной аналитики. – URL: https://webiomed.ru/blog/osnovnye-metriki-zadach-klassifikatsii-v-mashinnom-obuchenii/ (дата обращения 24.05.2023).

69. Первый международный стандарт в области промышленного интернета вещей утверждён на основе российских разработок [Электронный ресурс] // Росстандарт. – URL: https://www.gost.ru/portal/gost/home/presscenter/news/newsRST/?portal:isSecure=true&navigationalstate=JBPNS_rO0ABXczAAZhY3Rpb24AAAABAA5zaW5nbGVOZXdzVmlldwACaWQAAAABAAQ4NDg5AAdfX0VPRl9f&portal:componentId=88beae40-0e16-414c-b176-d0ab5de82e16 (дата обращения: 02.05.2023).

70. Петренко В.И., Тебуева Ф.Б., Павлов А.С., Стручков И.В. Анализ рисков нарушения информационной безопасности в роевых робототехнических системах при масштабировании численности агентов // Прикаспийский журнал: управление и высокие технологии. – 2022. – №2 (58). – С. 92-109.

71. Плаван А.И., Карташевский В.Г., Поздняк И.С. Сравнительный анализ статистических характеристик DDoS-атак и нормального трафика [Электронный ресурс] // Актуальные проблемы науки и техники: матер. I Междунар. науч.-техн. конф. (Сарапул, май 2021 г). – Ижевск : Изд-во УИР ИжГТУ имени М.Т. Калашникова, 2021. – 875 с.

72. Полтавцева М.А. Модель активного мониторинга как основа управления безопасностью промышленных киберфизических систем // Вопросы кибербезопасности. – 2021. – №2 (42). – С. 51-60.

73. Садков А. Алгоритмы маршрутизации и самоорганизации: лекция [Электронный ресурс] // Лаборатория физических основ и технологий беспроводной связи. Университет Лобачевского. – URL: https://wl.unn.ru/materials/courses/WSN/Lecture%204/Lecture4.ppt (дата обращения: 10.05.2023).

74. Сексембаева М.А. Особенности обеспечения безопасности в промышленном Интернете вещей [Электронный ресурс] // E-Scio. – 2019. – №5

(32). – URL: https://cyberleninka.ru/article/n/osobennosti-obespecheniya-bezopasnosti-v-promyshlennom-internete-veschey (дата обращения: 18.03.2023).

75. Сетевые аномалии. Что это и как их определить? [Электронный ресурс] // SecurityLab. – URL: https://www.securitylab.ru/analytics/535530.php (дата обращения: 21.01.2023).

76. Старовойтов В.В., Голуб Ю.И. Нормализация данных в машинном обучении // Информатика. – Т. 18. – № 3. – С.83-96.

77. Сухов В.Е. Система обнаружения аномалий сетевого трафика на основе искусственных иммунных систем и нейросетевых детекторов // Вестник РГРТУ. – 2015. – №54. – С.84-90.

78. Сухостат Л. Обнаружение атак на киберфизический системы на основе глубокого обучения [Электронный ресурс] // İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri: IV respublika konfransı, 14 dekabr 2018-ci il [Актуальные междисциплинарные научно-практические проблемы информационной безопасности: IV республиканская конференция, республика Азербайджан]. – С. 42-46. – URL: https://ict.az/uploads/konfrans/info_sec_2018/RS07_DETECTION-OF-ATTACKS-ON_CYBER-PHYSICAL-SYSTEMS-BASED-ON-DEEP-LEARNING.PDf (дата обращения: 30.04.2023).

79. Сычугов А.А. Информационная система оперативного обнаружения опасных состояний промышленных объектов // Известия ТулГУ. Технические науки. – 2021. – № 10. – С. 401-406.

80. Сычугов А.А., Греков М.М. Применение генеративных состязательных сетей в системах обнаружения аномалий [Электронный ресурс] // Моделирование, оптимизация и информационные технологии. – 2021. – № 9(1). – С.1-9. – URL: https://moitvivt.ru/ru/journal/article?id=921 (дата обращения 23.04.2022).

81. Татарникова Т.М., Богданов П.Ю. Обнаружение атак в сетях Интернета вещей методами машинного обучения // Информационно-управляющие системы. – 2021. – №6 (115). – С. 42-52.

82. Татарникова Т.М., Богданов П.Ю., Тимочкина Т.В. Комплексная защита Интернета вещей // Информационные системы и технологии в моделировании и управлении : сборник трудов VI Международной научно-практической конференции (24-26 мая 2021 г.) / К.А. Маковейчук. – Симферополь : АРИАЛ, 2021. – 414 с.

83. Тимочкина Т.В., Татарникова Т.М., Пойманова Е.Д. Применение нейронных сетей для обнаружения сетевых атак // Приборостроение. – 2021. – № 5. – С. 357-363.

84. Умная нормализация данных [Электронный ресурс] // Хабр. – URL: https://habr.com/ru/articles/527334/ (дата обращения 11.06.2023).

85. Филькин К.Н., Филькин С.Н., Шелупанов А.А. Информационно-управляющая система поддержки принятия решений при управлении информационной безопасностью территориально-распределенной организации // Безопасность информационных технологий. – 2007. – №4. – С.83-86.

86. Частикова В.А., Картамышев Д.А. Искусственные иммунные системы: основные подходы и особенности их реализации // Научные труды КубГТУ. – 2016. – № 8. – 193-208.

87. Чернов Д.В., Сычугов А.А. Современные подходы к обеспечению информационной безопасности АСУ ТП // Известия ТулГУ. Технические науки. – 2018. – №10. – С. 59-64.

88. Baseline security recommendations for IoT [Электронный ресурс] // ENISA. – URL: https://www.enisa.europa.eu/publications/baseline-security-recommendations-for-iot (дата обращения 29.05.2020).

89. Good practices for security of Internet of Things in the context of smart manufacturing [Электронный ресурс] // ENISA. – URL: https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot (дата обращения 29.05.2020).

90. Good practices for security of IoT – secure software development lifecycle [Электронный ресурс] // ENISA. URL: https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot-1 (дата обращения 29.05.2020).

91. ISO/IEC 27400:2022 Cybersecurity – IoT security and privacy – Guidelines [Электронный ресурс] // ISO. – URL: https://www.iso.org/standard/44373.html (дата обращения: 02.03.2021).

92. ISO/IEC 30147:2021 Information technology – Internet of things – Methodology for trustworthiness of IoT system/service [Электронный ресурс] // ISO. – URL: https://www.iso.org/standard/53267.html (дата обращения: 02.03.2021).

93. ISO/IEC TR 29181-5:2014 Information technology – Future network – Problem statement and requirements – Part 5: Security [Электронный ресурс] // ISO. – URL: https://www.iso.org/standard/57487.html (дата обращения: 02.03.2021).

94. NISTIR 8200. Interagency report on status of international cybersecurity standardization for the Internet of Things (IoT) / Interagency international cybersecurity standardization working group, 2018 [Электронный ресурс]. – URL: https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8200.pdf (дата обращения: 12.03.2021).

95. NISTIR 8228. Considerations for managing Internet of things (IoT) cybersecurity and privacy risks [Электронный ресурс] // Computer security resource center. NIST. – URL: https://csrc.nist.gov/publications/detail/nistir/8228/final (дата обращения: 02.03.2021).

96. Q.3913 : Set of parameters for monitoring Internet of things devices [Электронный ресурс] // ITU. – URL: https://www.itu.int/rec/T-REC-Q.3913/en (дата обращения: 02.03.2021).

97. Recommended practice: improving industrial control system cybersecurity with defense-in-depth strategies [Электронный ресурс] // Cybersecurity & infrastructure security agency. – URL: https://www.us-cert.gov/sites/default/files/recommended_practices/NCCIC_ICS-CERT_Defense_in_Depth_2016_S508C.pdf (дата обращения: 29.05.2020).

98. White Paper (Draft). Internet of things (IoT) trust concerns [Электронный ресурс] // Computer security resource center. NIST. – URL: https://csrc.nist.gov/publications/detail/white-paper/2018/10/17/iot-trust-concerns/draft (дата обращения: 02.03.2021).

99. Y.4455 : Reference architecture for Internet of things network capability exposure [Электронный ресурс] // ITU. – URL: https://www.itu.int/rec/T-REC-Y.4455-201710-P (дата обращения: 02.03.2021).

100. Y.4806 : Security capabilities supporting safety of the Internet of things [Электронный ресурс] // ITU. – URL: https://www.itu.int/rec/T-REC-Y.4806/en (дата обращения: 02.03.2021).

101. Abosata N., Al-Rubaye S., Inalhan G. Customised intrusion detection for an industrial IoT heterogeneous network based on machine learning algorithms called FTL-CID [Электронный ресурс] // Sensors. – 2023. – № 23. – URL: https://www.mdpi.com/1424-8220/23/1/321 (дата обращения 04.04.2023).

102. Alaparthy V., Morgera S. A multi-level intrusion detection system for wireless sensor networks based on immune theory // IEEE Access. – 2018. – №. 6. – С. 47364–47373.

103. Aldhaheri S., Alghazzawi D., Cheng L., Alzahrani B., Al-Barakat A. DeepDCA: novel network-based detection of IoT attacks using artificial immune system // Applied Sciences. – 2020. – №. 10. – С. 1909–1932.

104. Al-Duwairi B., Al-Kahla W., AlRefai M.A., Abdelqader Y., Rawash A., Fahmawi R. SIEM-based detection and mitigation of IoT-botnet DDoS attacks // International journal of electrical and computer engineering (IJECE). – 2020. – №2 (10). – С. 2182-2191.

105. Alem S., Espes D., Martin E., Nana L., Lamotte F. A hybrid intrusion detection system in industry 4.0 based on ISA95 standard [Электронный ресурс] // 2019 IEEE/ACS 16th international conference on Computer systems and applications (AICCSA). – 2019. – С.1-8. – URL: https://hal.archives-ouvertes.fr/hal-02506109v2/document (дата обращения: 30.07.2021).

106. Almomani I., Al-Kasasbeh B., AL-Akhras M. WSN-DS: a dataset for intrusion detection systems in wireless sensor networks [Электронный ресурс] // Journal of Sensors. – 2016. – vol. 2016. – URL: https://www.hindawi.com/journals/js/2016/4731953/ (дата обращения: 01.03.2020).

107. Almomani I., Kasasbeh B. Performance analysis of LEACH protocol under denial of service attacks [Электронный ресурс] // 2015 6th International conference on information and communication systems (ICICS), Researchgate. – URL: https://www.researchgate.net/publication/276288461_Performance_analysis_of_LEACH_protocol_under_Denial_of_Service_attacks (дата обращения 14.03.2023).

108. Almseidin M., Alkasassbeh M. An accurate detection approach for IoT botnet attacks using interpolation reasoning method [Электронный ресурс] // Information. – 2022. – Т. 13. – № 6. – URL: https://www.mdpi.com/2078-2489/13/6/300 (дата обращения: 24.05.2023).

109. Alqahtani M., Gumaei A., Mathkour H., Ismail M.M.B. A genetic-based extreme gradient boosting model for detecting intrusions in wireless sensor networks // Sensors. – 2019. – №. 19. – С. 1-20.

110. Alsulaimanand L., Al-Ahmadi S. Performance evaluation of machine learning techniques for DoS detection in wireless sensor network // International journal of network security & its applications (IJNSA). – 2021. – №. 2 (13). – С. 21-29.

111. Ammar M., Russello G., Crispo B. Internet of things: a survey on the security of IoT frameworks [Электронный ресурс] // Journal of information security and applications. – 2018. – №38. – С. 8-27. – URL: https://www.sciencedirect.com/science/article/abs/pii/S2214212617302934?via (дата обращения: 02.08.2022).

112. Ankita, Rani S., Singh A., Elkamchouchi D.H., Noya I.D. Lightweight hybrid deep learning architecture and model for security in IIoT [Электронный ресурс] // Applied sciences. – 2022. – № 12 (13). URL: https://www.mdpi.com/2076-3417/12/13/6442 (дата обращения 04.04.2023).

113. Aposemat IoT-23 [Электронный ресурс] // Stratosphere Lab. – URL: https://www.stratosphereips.org/datasets-iot23 (дата обращения: 05.04.2023).

114. Argus [Электронный ресурс] // Openargus. – URL: https://openargus.org/ (дата обращения 14.03.2023).

115. AWID2 [Электронный ресурс] // University of the Aegean. – URL: https://icsdweb.aegean.gr/awid/awid2 (дата обращения: 30.03.2023).

116. AWID3 [Электронный ресурс] // University of the Aegean. – URL: https://icsdweb.aegean.gr/awid/awid3 (дата обращения: 30.03.2023).

117. Bahaa A., Sayed A., Elfangary L., Fahmy H. A novel hybrid optimization enabled robust CNN algorithm for an IoT network intrusion detection approach [Электронный ресурс] // PLoS ONE. – 2022. – № 17 (12). – URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0278493 (дата обращения 04.04.2023).

118. CAN Dataset for intrusion detection (OTIDS) [Электронный ресурс] // HCRL. – URL: https://ocslab.hksecurity.net/Dataset/CAN-intrusion-dataset (дата обращения: 05.04.2023).

119. Chavez A., Lai C., Jacobs N., Hossain-McKenzie S., Jones C.B., Johnson J., Summers A. Hybrid intrusion detection system design for distributed energy resource systems [Электронный ресурс] // IEEE CyberPELS. – 2019. – URL: https://ieeexplore.ieee.org/document/8925064 (дата обращения: 30.07.2021).

120. Check Point IoT Protect [Электронный ресурс] // CheckPoint. – URL: https://www.checkpoint.com/downloads/products/cp-iot-security-solution-brief.pdf (дата обращения: 13.03.2021).

121. Chou T.S., Yen K.K., Luo J. Network intrusion detection design using feature selection of soft computing paradigms // International journal of computer and information engineering. – 2008. – Т. 2. – № 11. – С. 3722-3734.

122. Combining multiple learners: lecture notes for E. Alpaydın 2004 Introduction to Machine Learning [Электронный ресурс] / The MIT Press (V1.1) //

Sabanci Universitesi – URL: http://people.sabanciuniv.edu/berrin/cs512/lectures/9-i2ml-chap15-classifier-combination-short.pdf (дата обращения: 24.09.2021).

123. Damasevicius R., Venckauskas A., Grigaliunas S., Toldinas J., Morkevicius N., Aleliunas T., Smuikys P. LITNET-2020: an annotated real-world network flow dataset for network intrusion detection [Электронный ресурс] // Electronics. – 2020. – № 9. – С. 1-23. – URL: https://www.mdpi.com/2079-9292/9/5/800 (дата обращения: 29.03.2023).

124. Dataset of legitimate IoT data VARIoT [Электронный ресурс] // data.gouv.fr : Open platform for French public data. – URL: https://www.data.gouv.fr/en/datasets/dataset-of-legitimate-iot-data/ (дата обращения: 05.04.2023).

125. Dilek S., Çakır H., Aydın M. Applications of artificial intelligence techniques to combating cyber crimes: a review // International journal of artificial intelligence & applications (IJAIA). – 2015. – Т. 6. – № 1. – С. 21-39.

126. Dong R.-H., Yan H.-H., Zhang Q.-Y. An intrusion detection model for wireless sensor network based on information gain ratio and bagging algorithm // International journal of network security. – 2020. – №. 2 (22). – С. 218-230.

127. DS2OS traffic traces [Электронный ресурс] // Kaggle. – URL: https://www.kaggle.com/datasets/francoisxa/ds2ostraffictraces (дата обращения: 05.04.2023).

128. Dutta V., Choras M., Pawlicki M., Kozik R. A deep learning ensemble for network anomaly and cyber-attack detection [Электронный ресурс] // Sensor. – 2020. – № 20 (16). – С. 1-20. – URL: https://www.mdpi.com/1424-8220/20/16/4583 (дата обращения 04.04.2023).

129. Estlund D.M. Opinion leaders, independence, and Condorcet's jury theorem // Theory and decision. – 1994. – Т. 36. – С. 131-162.

130. Farooq N., Zahoor I., Mandal S., Gulzar T. Systematic analysis of DoS attacks in wireless sensor networks with wormhole injection // International journal of information and computation technology. – 2014. – №2 (4). – С. 173–182.

131. Garcia S., Parmisano A., Erquiaga M.J. IoT-23: a labeled dataset with malicious and benign IoT network traffic (Version 1.0.0) [Электронный ресурс] // Zenodo. – URL: http://doi.org/10.5281/zenodo.4743746 (дата обращения: 05.04.2023).

132. Golchha R., Joshi A., Gupta G.P. Voting-based ensemble learning approach for cyber attacks detection in industrial Internet of things [Электронный ресурс] // Procedia Computer science. – 2023. – № 218. – 1752-1759. – URL: https://www.sciencedirect.com/science/article/pii/S1877050923001539?via (дата обращения 15.05.2023).

133. Golovko V., Komar M., Sachenko A. Principles of neural network artificial immune system design to detect attacks on computers [Электронный ресурс] // International conference on Modern problems of radio engineering, telecommunications and computer science (TCSET). – 2010. – С. 237. – URL: https://ieeexplore.ieee.org/document/5446089 (дата обращения: 30.07.2021).

134. Gopali S., Namin A.S. Deep learning-based time-series analysis for detecting anomalies in Internet of Things [Электронный ресурс] // Electronics. – 2022. – № 11. – URL: https://www.mdpi.com/2079-9292/11/19/3205 (дата обращения 04.04.2023).

135. Hussein M.A., Hamza E.K. Secure mechanism applied to big data for IIoT by using security event and information management system (SIEM) // International journal of intelligent engineering and systems. – 2022. – №6 (15). – С. 667-681. – URL: https://inass.org/wp-content/uploads/2022/09/2022123159-2.pdf (дата обращения 21.03.2023).

136. ICT219 Lecture 11 – Hybrid intelligent systems [Электронный ресурс] // StuDocu. – URL: https://www.studocu.com/en-au/document/murdoch-university/intelligent-systems/ict219-lecture-11-hybrid-intelligent-systems/1280311 (дата обращения: 30.07.2021).

137. Igbe O., Darwish I., Saadawi T. Distributed network intrusion detection system: an artificial immune system approach // 2016 IEEE first international conference on Connected health: applications, systems and engineering technologies (CHASE), Washington D.C., the USA. – С. 101-106.

138. IoT dataset for intrusion detection systems (IDS) [Электронный ресурс] // Kaggle. – URL: https://www.kaggle.com/datasets/azalhowaide/iot-dataset-for-intrusion-detection-systems-ids (дата обращения: 05.04.2023).

139. IoT или АСУ ТП? [Электронный ресурс] // Диплайн. – URL: http://diplinegroup.ru/novosti/iot-ili-asu-tp.html (дата обращения: 23.05.2023).

140. Kali Linux [Электронный ресурс] // Kali. – URL: https://www.kali.org/ (дата обращения 15.05.2023)

141. Karabiber F. Precision and Recall [Электронный ресурс] // Learndatasci. – URL: https://www.learndatasci.com/glossary/precision-and-recall/ (дата обращения 24.05.2023).

142. KDD Cup 1999 Data [Электронный ресурс]. – URL: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (дата обращения: 05.02.2018).

143. Khan M.A., Kim Y. Deep learning-based hybrid intelligent intrusion detection system // Computers, materials & continua. – 2021. – №1 (68). – С. 671-687.

144. Komar M., Golovko V., Sachenko A., Bezobrazov S. Development of neural network immune detectors for computer attacks recognition and classification // 2013 IEEE 7th international conference on Intelligent data acquisition and advanced computing systems (IDAACS). – 2013. – С. 665-668.

145. Komar M., Sachenko A., Bezobrazov S., Golovko V. Intelligent cyber defense system using artificial neural network and immune system techniques // Information and communication technologies in education, Research and industrial applications. ICTERI 2016. Communications in Computer and Information science. – Т. 783. – URL: https://link.springer.com/chapter/10.1007/978-3-319-69965-3_3 (дата обращения: 18.05.2023).

146. Kotenko I., Fedorchenko A., Saenko I., Kushnerevich A. Big data technologies for security event correlation based on event type accounting // Вопросы кибербезопасности. – 2017. – № 5 (24). – С. 2-16.

147. Kumar P., Gupta G.P., Tripathi R. An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks // Computer communications. – 2021. – № 166. – С.110-124. – URL: https://www.sciencedirect.com/science/article/abs/pii/S0140366420320090?via (дата обращения: 30.04.2023).

148. Kumaravel H.V. An anomaly-based intrusion detection system based on artificial immune system (AIS) techniques. Open access theses. 2016:964. Доступно по: https://docs.lib.purdue.edu/open_access_theses/964 (дата обращения: 25.08.2022).

149. Latif S., Zou Z., Idrees Z., Ahmad J. A novel attack detection scheme for the industrial Internet of things using a lightweight random neural network [Электронный ресурс] // EEE Access. – 2020. – №. 8. – С. 89337-89350. – URL: https://ieeexplore.ieee.org/abstract/document/9091574 (дата обращения: 05.06.2022).

150. Le T.-T.-H., Park T., Cho D., Kim H. An effective classification for DoS attacks in wireless sensor networks [Электронный ресурс] // 2018 10th international conference on Ubiquitous and future networks (ICUFN). – 2018. – С. 689–692. – URL: https://www.researchgate.net/publication/327065277_An_Effective_Classific ation_for_DoS_Attacks_in_Wireless_Sensor_Networks (дата обращения: 06.09.2021).

151. Li Y., Jing C., Xu J. A new distributed intrusion detection method based on immune mobile agent // Life system modeling and intelligent computing. ICSEE 2010, LSMS 2010, Lecture notes in computer science. – №. 6328. – С. 233-243.

152. Machine learning-based NIDS datasets [Электронный ресурс] // the University of Queensland. – URL: https://staff.itee.uq.edu.au/marius/NIDS_datasets/ (дата обращения: 05.04.2023).

153. Mahboubian M., Hamid N.A.W.A. A machine learning based AIS IDS // International journal of machine learning and computing. – 2013. – №3 (3). – С. 259-262.

154. Meidan Y., Bohadana M., Mathov Y., Mirsky Y., Breitenbacher D., Shabtai A., Elovici Y. N-BaIoT: network-based detection of IoT botnet attacks using deep autoencoders [Электронный ресурс] // IEEE pervasive computing. – 2018. – № 3. – Т. 17. – С. 12-22. – URL: https://ieeexplore.ieee.org/document/8490192 (дата обращения: 05.04.2023).

155. Mitrokotsa A., Karygiannis T. Intrusion detection techniques in sensor networks // Wireless sensor network security, Cryptology and information security series. IOS Press, 2008. – С. 251–272.

156. Ms. Saranya V.S., Dr. Ramachandran G., Dr. Chakaravarthi S. An intelligent IoT attack detection framework using effective edge AI based computing // Indian journal of computer science and engineering. – 2022. – № 4 (13). – С. 1156-1167.

157. Mukkamala S., Sung A.H. Identifying significant features for network forensic analysis using artificial intelligent techniques // International journal of digital evidence. – 2003. – №. 1 (4). – С. 1-17.

158. Mukkamala S., Sung A.H., Abraham A. Modeling intrusion detection systems using linear genetic programming approach. [Электронный ресурс]. – URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.83.544&rep=rep1&type=pdf (дата обращения: 29.02.2018).

159. Nespoli P., Marmol F.G. E-Health wireless IDS with SIEM integration [Электронный ресурс] // Conference: WCNC'18, Barcelona, 2018. – URL: https://www.researchgate.net/publication/324482808_e-Health_Wireless_IDS_with_SIEM_integration (дата обращения 21.02.2023).

160. NF-BoT-IoT [Электронный ресурс] // Kaggle. – URL: https://www.kaggle.com/datasets/dhoogla/nfbotiot (дата обращения: 05.04.2023).

161. NSL-KDD dataset [Электронный ресурс] // University of New Brunswick. – URL: https://www.unb.ca/cic/datasets/nsl.html (дата обращения: 05.06.2022).

162. Nugroho E.P., Djatna T., Sitanggang I.S., Buono A., Hermadi I. A review of intrusion detection system in IoT with machine learning approach: current and future research // 2020 6th International conference on science in information technology (ICSITech), Палу, Индонезия, 2020. – С. 138-143.

163. Oates R., Greensmith J., Aickelin U. The application of a dendritic cell algorithm to a robotic classifier // Proceedings of ICARIS'07. Берлин: Springer, 2007. – С. 204-215.

164. OSSIM – разворачиваем комплексную open source систему управления безопасностью [Электронный ресурс] // Хабр. – URL: https://habr.com/ru/articles/255433/ (дата обращения: 25.05.2023).

165. OSSIM [Электронный ресурс]. – URL: https://cybersecurity.att.com/products/ossim (дата обращения 25.04.2023)

166. Ovasapyan T., Moskvin D. Security provision in WSN on the basis of the adaptive behavior of nodes [Электронный ресурс] // Proceedings of the 4th world conference on Smart trends in systems, security and sustainability (WorldS4). – URL: https://ieeexplore.ieee.org/document/9210421 (дата обращения: 05.03.2021).

167. Panda M., Abraham A. Patra M.R. A hybrid intelligent approach for network intrusion detection // Procedia Engineering. – 2012. – №30. – С.1-9.

168. Panda M., Abraham A., Patra M.R. Hybrid intelligent systems for detecting network intrusions [Электронный ресурс] // Security and communication networks. – 2012. – №8 (16). – URL: https://www.researchgate.net/publication/260408971_Hybrid_intelligent_systems_for_detecting_network_intrusions (дата обращения: 30.07.2021).

169. Powers S.T., He J. A hybrid artificial immune system and self organising map for network intrusion detection // Information Sciences. – 2008. – № 15(178). – С. 3024-3042.

170. Raja K., Karthikeyan K., Abilash B, Dev K., Raja G. Deep learning based attack detection in IIoT using two-level intrusion detection system: предварительная версия статьи [Электронный ресурс] // Research square, 2021. – URL: https://www.researchsquare.com/article/rs-997888/v1 (дата обращения 04.04.2023).

171. Sagu A., Gill N.S., Gulia P., Chatterjee J.M., Priyadarshini I. A hybrid deep learning model with self-improved optimization algorithm for detection of security attacks in IoT environment [Электронный ресурс] // Future internet. – 2022. – № 14 (10). – URL: https://www.mdpi.com/1999-5903/14/10/301 (дата обращения 04.04.2023).

172. Salama M.A., Ramadan R., Darwish A., Eid H.F. Hybrid intelligent intrusion detection scheme // Advances in intelligent and soft computing. – 2011. – № 96. – С. 295-302.

173. Sarhan M., Layeghy S., Moustafa N., Portmann M. NetFlow datasets for machine learning-based network intrusion detection systems [Электронный ресурс] // Big data technologies and applications. BDTA WiCON 2020 : lecture notes of the Institute for Computer sciences, Social informatics and Telecommunications engineering. – 2020. – № 371. Springer, Cham. – URL: https://doi.org/10.1007/978-3-030-72802-1_9 (дата обращения: 10.04.2023).

174. Sen J. Security in wireless sensor networks // Wireless sensor networks: current status and future trends. – New York: CRC Press, 2012. – С. 407-460.

175. Sharabyrov I., Vasilyev V., Guzairov M., Mashkina I. Wireless intrusion detection system on the basis of data mining methods // Proceedings of the 13th International conference on applied computing, oct. 28-30, 2016, Мангейм, Германия, 2016. – С. 43-50.

176. Sharma A., Sharma D. Clonal selection algorithm for classification / P. Liò, G. Nicosia, T. Stibor // Artificial immune systems. ICARIS 2011 : lecture notes in Computer Science. – Берлин, Гейдельберг: Springer, 2011. – Т. 6825. – URL: https://doi.org/10.1007/978-3-642-22371-6_31 (дата обращения: 29.04.2023).

177. SIEM системы: найти иголку в стогу сена [Электронный ресурс] // ИнфоБезпека. – URL: http://www.infobezpeka.com/publications/?id=589 (Дата обращения: 08.12.2022).

178. Snort [Электронный ресурс]. – URL: https://www.snort.org/ (дата обращения: 25.05.2023).

179. Tarakanov A.O., Tarakanov Y.A. A comparison of immune and genetic algorithms for two real-life tasks of pattern recognition // Int. J. of unconventional computing. – 2004. – Т. 1.4. – С. 357-374.

180. Tarakanov A.O., Tarakanov Y.A. A comparison of immune and neural computing for two real-life tasks of pattern recognition // International conference on Artificial immune systems, Catania, 2004. – С. 236-249.

181. Teerawat I., Hossain E. Introduction to network simulator NS2 [Электронный ресурс] // Springer, 2011. – URL: https://link.springer.com/chapter/10.1007/978-1-4614-1406-3_2 (дата обращения 14.03.2023.

182. The Bot-IoT dataset [Электронный ресурс] // UNSW Sydney. – URL: https://research.unsw.edu.au/projects/bot-iot-dataset (дата обращения: 04.06.2022).

183. The ToN_IoT Dataset [Электронный ресурс] // UNSW Sydney. – URL: https://research.unsw.edu.au/projects/toniot-datasets (дата обращения: 05.04.2023).

184. The UNSW-NB15 Dataset [Электронный ресурс] // UNSW Sydney. – URL: https://research.unsw.edu.au/projects/unsw-nb15-dataset (дата обращения: 05.04.2023).

185. Threat intelligence report 2020 [Электронный ресурс] // NOKIA. – URL: https://pages.nokia.com/T005JU-Threat-Intelligence-Report-2020.html?_ga=2.216248470.1653315497.1608038999-829562352.1608038999 (дата обращения: 23.09.2021).

186. Tizio G.D., Massacci F., Allodi L., Dashevskyi S., Mirkovic J. An experimental approach for estimating cyber risk: a proposal building upon cyber ranges and capture the flags // 2020 IEEE European symposium on security and privacy workshops (EuroS&PW), Генуя, Италия, 2020. – С. 56-65.

187. Ullah I., Mahmoud Q.H. A scheme for generating a dataset for anomalous activity detection in IoT networks // Advances in artificial intelligence. Canadian AI 2020. Lecture notes in computer science. Springer, Cham. – 2020. – Т. 12109. – С.508-520.

188. Vaitsekhovich L. Intrusion detection in TCP/IP networks using immune systems paradigm and neural network detectors [Электронный ресурс] // XI International PhD Workshop OWD. – 2009. – С.219-224. – URL: https://www.researchgate.net/publication/306194779_Intrusion_detection_in_TCPIP_networks_using_immune_systems_paradigm_and_neural_network_detectors (дата обращения: 30.07.2021).

189. Vasilyev V., Shamsutdinov R. Distributed intelligent system of network traffic anomaly detection based on artificial immune system // Proceedings of the 7th scientific conference on information technologies for intelligent decision making support (ITIDS 2019), 28-30 may 2019., Ufa. Advances in intelligent system research. – 2019. – Т. 166. – С. 40-45.

190. Vasilyev V., Shamsutdinov R. Providing information security on the base of artificial immune system for industrial Internet of things: Proceedings of the 8th scientific conference on Information technologies for intelligent decision making support (ITIDS'2020) // Advances in intelligent systems research. – 2020. – Т. 174. – С. 212-217.

191. Vasilyev V., Shamsutdinov R. Security analysis of wireless sensor networks using SIEM and multi-agent approach [Электронный ресурс] // Proceedings of the Global smart industry conference (GloSIC'2020), 17-19 November 2020. – URL: https://ieeexplore.ieee.org/document/9267830 (дата обращения: 05.06.2022).

192. Wang L., Jajodia S., Singhal A., Noel S. k-Zero Day Safety: Measuring the security risk of networks against unknown attacks [Электронный ресурс] // Computer security resource center. – URL: https://csrc.nist.gov/CSRC/media/Projects/Measuring-Security-Risk-in-Enterprise-Networks/documents/k0day_august-version.pdf (дата обращения 03.02.2023).

193. Wireshark [Электронный ресурс] // Wireshark. – URL: https://www.wireshark.org/ (дата обращения 14.03.2023).

194. WUSTL-IIOT-2021 dataset for IIoT cybersecurity research [Электронный ресурс] // Washington University in St. Louis. – URL: https://www.cse.wustl.edu/~jain/iiot2/index.html (дата обращения 24.04.2023).

195. Xiao X., Zhang R. A danger theory inspired protection approach for hierarchical wireless sensor networks // KSII Transactions on Internet and information systems. – 2019. – Т. 13. – №. 5. – С. 2732–2753.

196. Yang S., Shiue Y., Su Z., Liu I., Liu C. An authentication information exchange scheme in WSN for IoT applications // IEEE Access. – 2020. – № 8. – С. 9728–9738

197. Yu K., Tan L., Mumtaz S., Al-Rubaye S., Al-Dulaimi A., Bashir A.K., Khan F.A. Securing critical infrastructures: deep learning-based threat detection in the IIoT [Электронный ресурс] // IEEE Communications Magazine. – 2021. – № 59 (10). – С. 76-82. – URL: https://e-space.mmu.ac.uk/631059/ (дата обращения 04.04.2023).

198. Zaind A., Maarof M., Shamsnddin S., Abraham A. Ensamble of one-class classifier for network intrusion detections [Электронный ресурс] // softcomputing.net. – URL: http://www.softcomputing.net/ias08_1.pdf (дата обращения: 29.02.2018).

199. Zeek [Электронный ресурс]. – URL: https://zeek.org/(дата обращения: 25.05.2023).

200. Zolanvari M., Teixeira M.A., Gupta L., Jain R. Machine learning based network vulnerability analysis of industrial Internet of Things // IEEE Internet of Things journal. – 2019. – Т. 6. – № 4. – С. 6822-2834. – URL: https://www.cse.wustl.edu/~jain/papers/vulnerab.htm (дата обращения 14.03.2023).

201. Zolanvari M., Teixeira M.A., Jain R. Effect of imbalanced datasets on security of industrial IoT using machine learning [Электронный ресурс] // 2018 IEEE international conference on intelligence and security informatics (ISI). – Майами, 2018. – URL: https://www.cse.wustl.edu/~jain/papers/ftp/imb_isi.pdf (дата обращения 14.03.2023).

Приложение А. Акты внедрения результатов работы

«УТВЕРЖДАЮ»
Директор
ЗАО «Республиканский центр
защиты информации»

_____С.Н. Зарипов
«25» апреля 2023 г.

АКТ
О внедрении ЗАО «Республиканский центр защиты информации (РЦЗИ)» результатов диссертационной работы Шамсутдинова Рината Рустемовича «Интеллектуальная система мониторинга информационной безопасности промышленного Интернета вещей с использованием механизмов искусственных иммунных систем»,
представленной на соискание ученой степени кандидата технических наук

Комиссия в составе: главный инженер ЗАО «РЦЗИ», к.т.н. Бакиров А.А.; ведущий специалист ЗАО «РЦЗИ» Федотов Д.Б.; ведущий специалист ЗАО «РЦЗИ» Кухарев С.Н. составили настоящий акт о том, что научно-технические результаты диссертационной работы Р.Р. Шамсутдинова «Интеллектуальная система мониторинга информационной безопасности промышленного Интернета вещей с использованием механизмов искусственных иммунных систем», представленной на соискание ученой степени кандидата технических наук:

– алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием адаптивных механизмов искусственных иммунных систем;

– алгоритмы обнаружения компьютерных атак и аномалий сетевого трафика промышленного Интернета вещей с использованием ансамбля методов машинного обучения и подсистемы корреляции событий ИБ;

– архитектура и методика применения гибридной распределенной интеллектуальной системы мониторинга ИБ промышленного Интернета вещей

используются в проектной работе ЗАО «РЦЗИ» на этапе анализа угроз безопасности информации и оценки рисков информационной безопасности, вследствие возможной реализации сетевых атак на промышленные системы.

Применение предложенных в диссертации алгоритмов позволяет повысить точность и оперативность обнаружения компьютерных атак на промышленные автоматизированные системы, без существенного увеличения требуемых вычислительных ресурсов.

Главный инженер ЗАО «РЦЗИ»,
к.т.н., _____А.А. Бакиров

Ведущий специалист ЗАО «РЦЗИ» _____Д.Б. Федотов

Ведущий специалист ЗАО «РЦЗИ» _____С.Н. Кухарев

«УТВЕРЖДАЮ»
Проректор по учебной работе
ФГБОУ ВО «Уфимский университет науки и технологий»
д-р физ.-мат. наук, профессор
_____Ю.В. Рахманова
«17» ____июля_____ 2023 г.

АКТ
О внедрении результатов диссертационной работы
Шамсутдинова Рината Рустемовича на тему:
«Интеллектуальная система мониторинга информационной безопасности промышленного Интернета вещей с использованием механизмов искусственных иммунных систем»,
представленной на соискание ученой степени кандидата технических наук

Комиссия в составе: заведующий кафедрой вычислительной техники и защиты информации (ВТиЗИ), д.ф.-м.н., профессор Картак В.М.; профессор кафедры ВТиЗИ, д.т.н., профессор Фрид А.И.; начальник Учебного управления, к.э.н., доцент Гумерова З.Ж., составила настоящий акт о том, что следующие результаты диссертационной работы Шамсутдинова Р.Р. используются в учебном процессе кафедры вычислительной техники и защиты информации:

– алгоритмы и методика применения гибридной распределенной интеллектуальной системы мониторинга информационной безопасности промышленного Интернета вещей.

Материалы диссертационной работы используются в лекционных курсах, а также при проведении практических и лабораторных занятий по дисциплинам «Искусственный интеллект в системах защиты информации», «Экспертные системы комплексной оценки безопасности информационно-телекоммуникационных систем» для обучающихся по направлениям подготовки специалистов 10.05.05 «Безопасность информационных технологий в правоохранительной сфере», магистров 10.04.01 «Информационная безопасность», 09.04.01 «Информатика и вычислительная техника» (профиль «Безопасность и защита информации»).

Результаты диссертационного исследования активно используются в учебном процессе при выполнении разделов курсовых проектов, научно-исследовательских и выпускных квалификационных работ, связанных с интеллектуальным обнаружением сетевых атак и оценкой рисков информационной безопасности промышленных объектов, систем промышленного Интернета вещей.

Заведующий кафедрой ВТиЗИ,
д.ф.-м.н., профессор _____ В.М. Картак

Профессор кафедры ВТиЗИ,
д.т.н., профессор _____ А.И. Фрид

Начальник учебного управления,
к.э.н., доцент _____ З.Ж. Гумерова

Приложение Б. Блок-схема алгоритма нормализации параметров сетевых соединений

Блок-схема предлагаемого алгоритма нормализации параметров ДС представлена на Рисунках Б.1-Б.2.

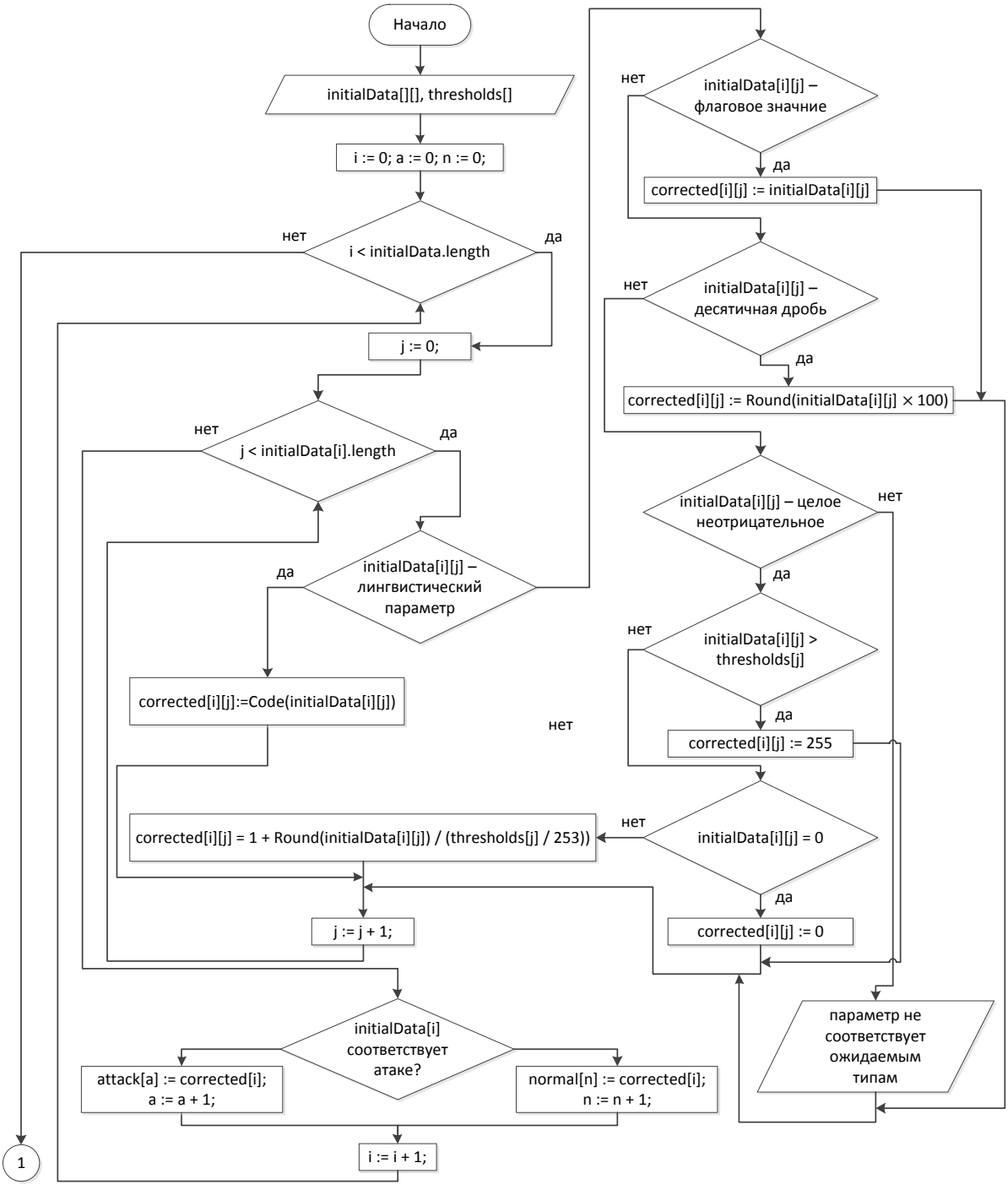


Рисунок Б.1 – Блок-схема нормализации параметров ДС. Часть 1

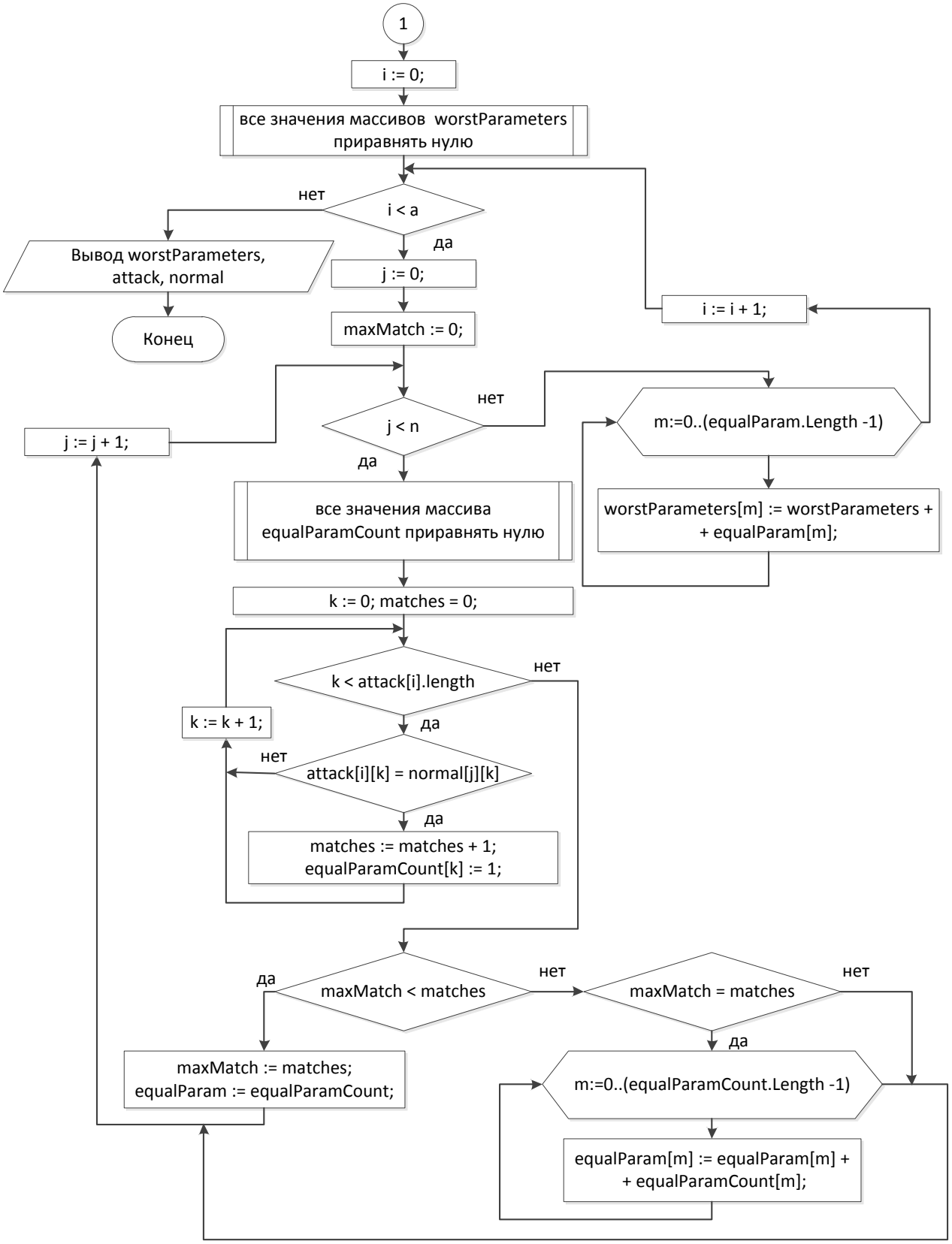


Рисунок Б.2 – Блок-схема нормализации параметров ДС. Часть 2

На основе выводимого массива *worstParameters*, являющегося вектором-строкой с количеством совпадений по каждому номеру параметра в соответствии с номером в векторе-строке, вычисляется частота совпадений параметров. Множества *attack* и *normal* представляют собой нормализованные данные об атаках и нормальном состоянии системы. *Round*() – функция округления до ближайшего целого, *Code*() – функция кодирования лингвистической переменной натуральным числом. В случае перехвата или получения реального сетевого трафика, нормализация выполняется в соответствии с блок-схемой, представленной на Рисунке Б.3.

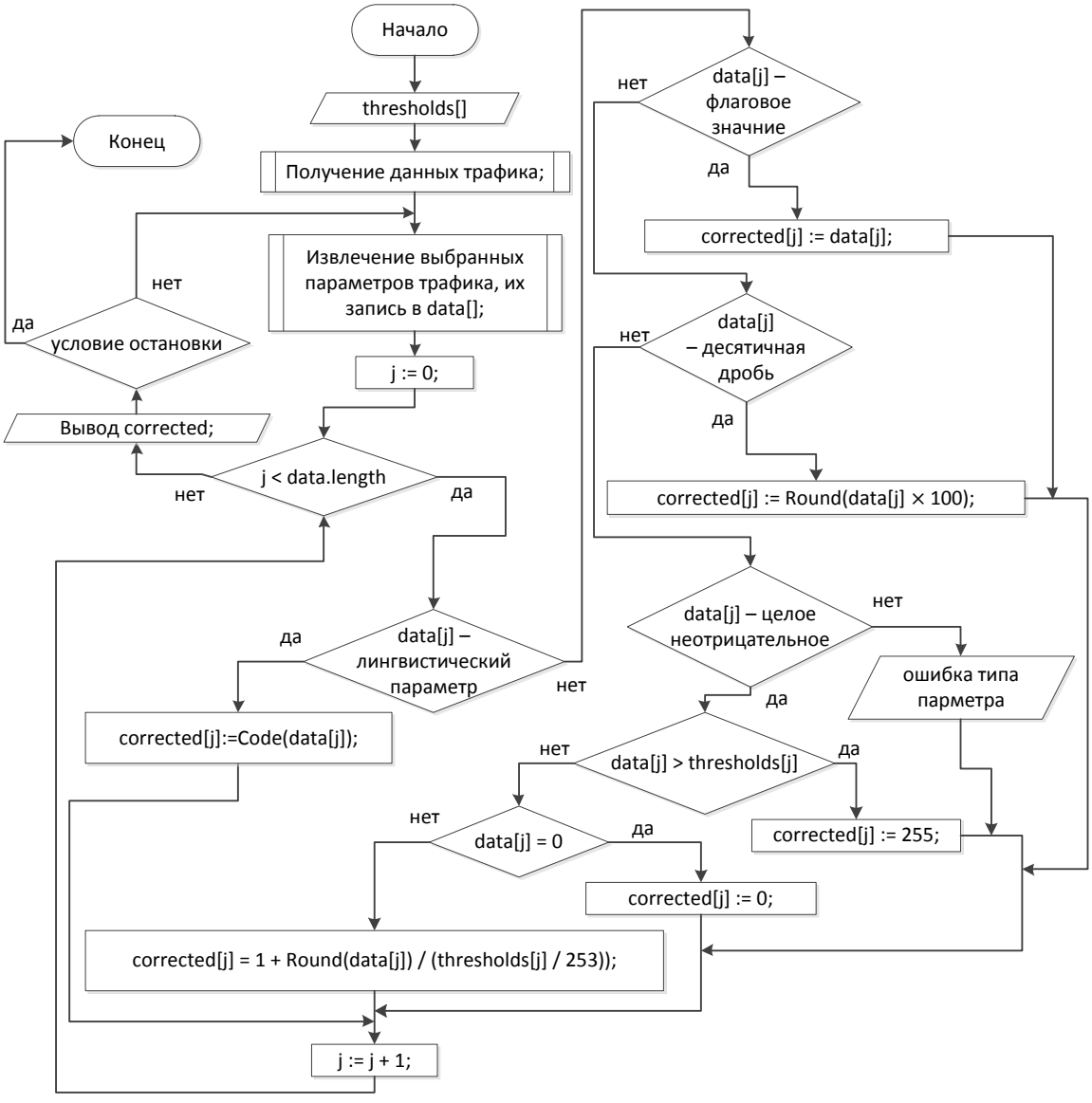


Рисунок Б.3 – Блок-схема нормализации потока реальных сетевых данных

Приложение В. Отчет подсистем дендритных клеток

Примеры отчетов, выводимых подсистемой ДК представлена Рисунками В.1-В.3.

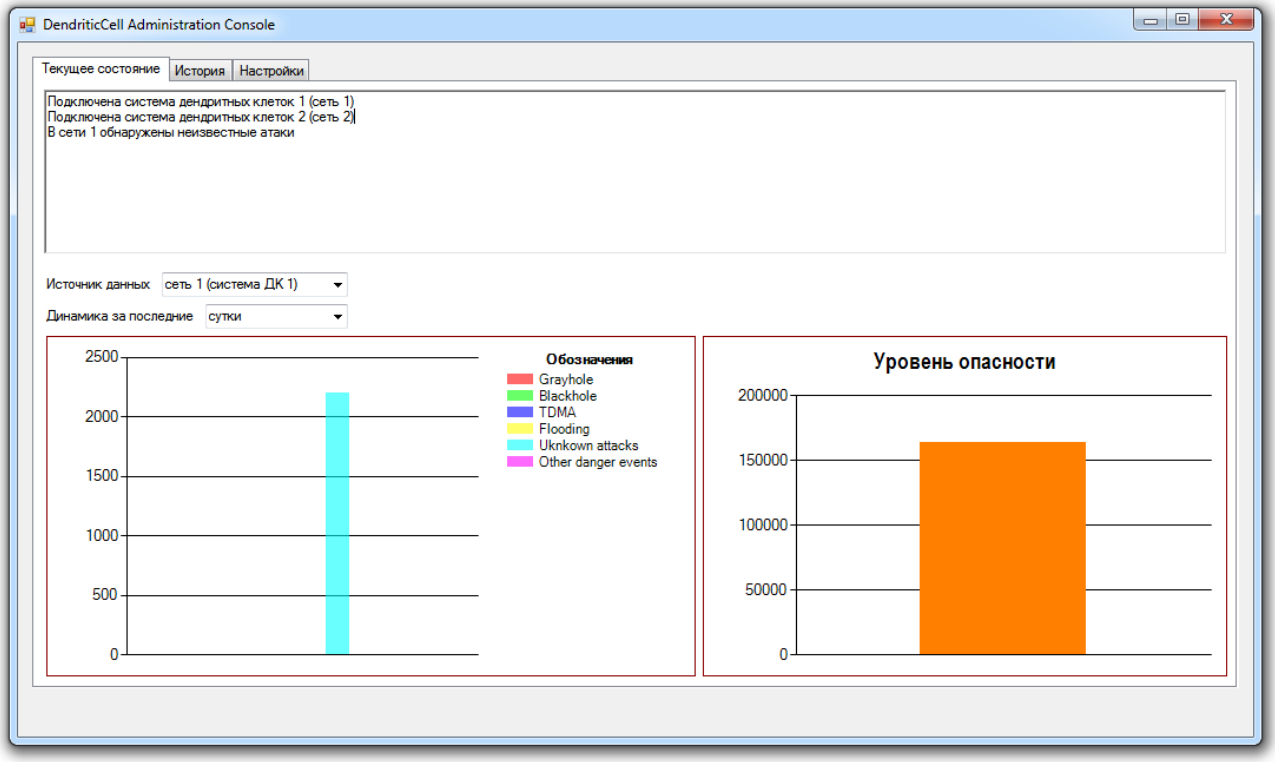


Рисунок В.1 – Отчет первой подсистемы ДК после двух эпох анализа.

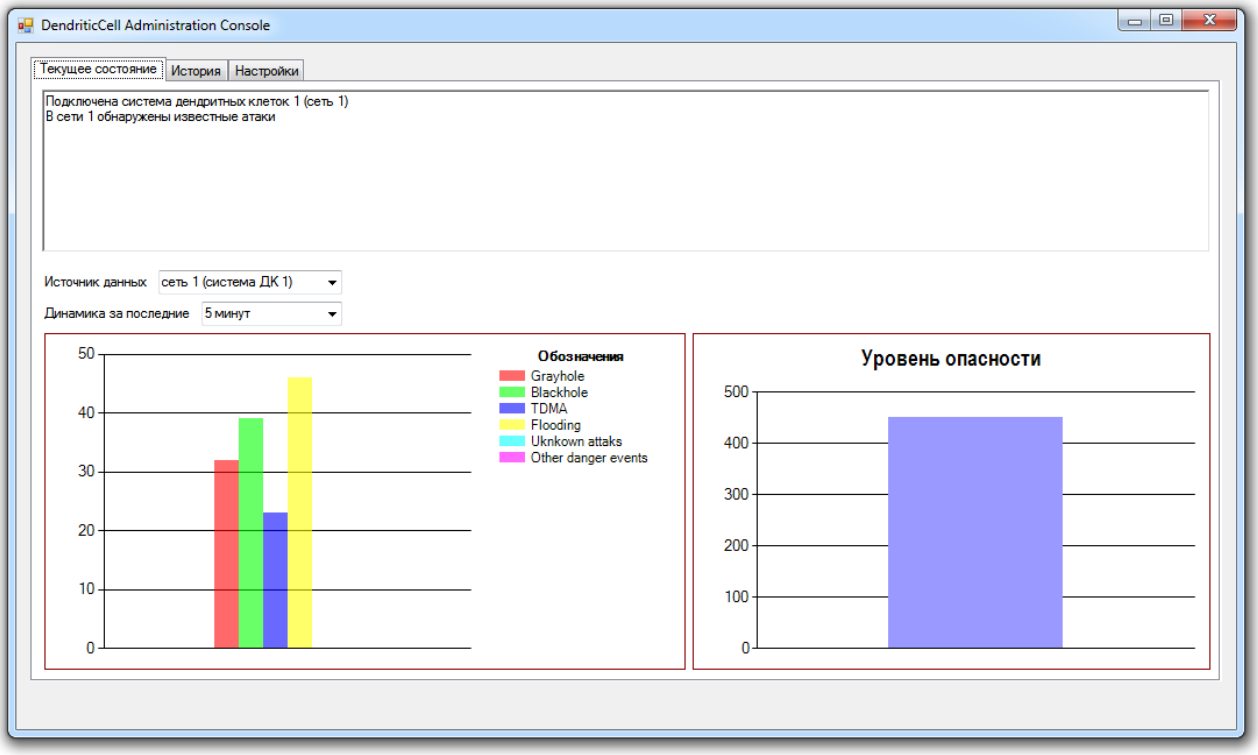


Рисунок В.2– Отчет первой подсистемы дендритных клеток с распознаванием атак

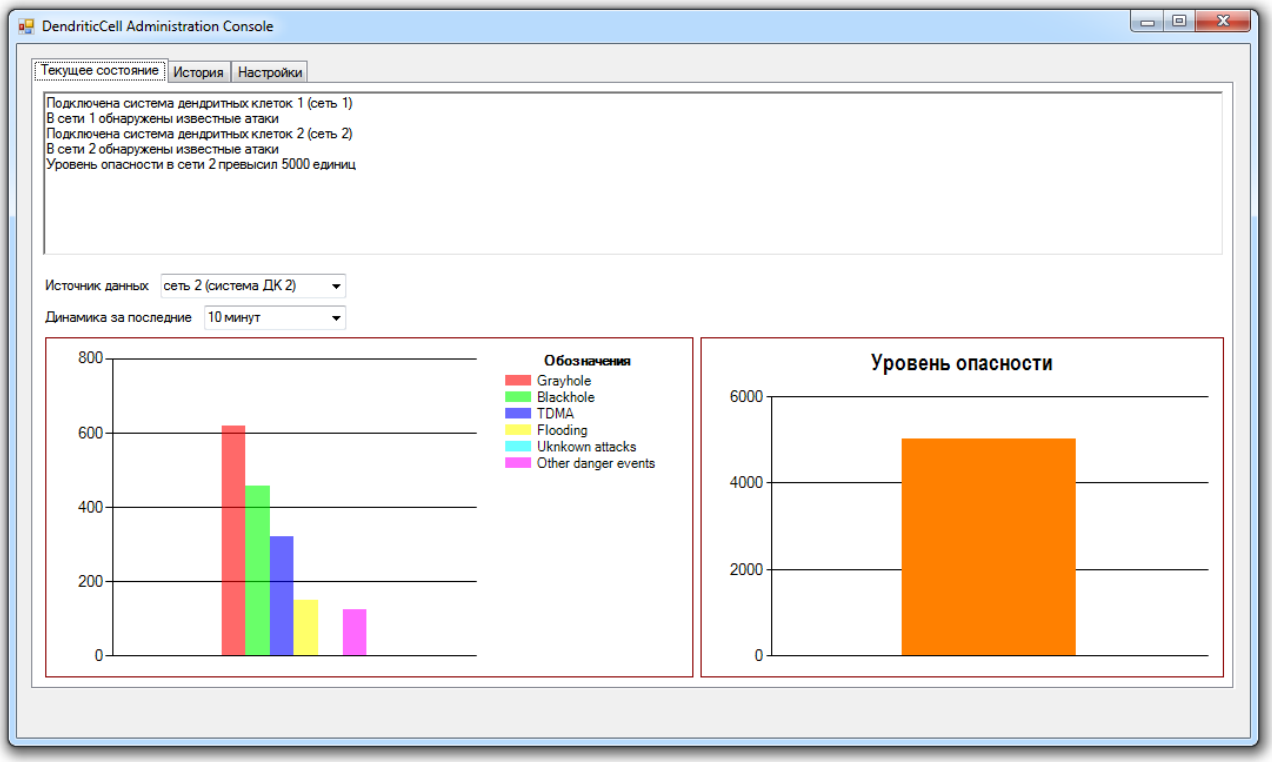


Рисунок В.3– Отчет второй подсистемы дендритных клеток с распознаванием атак