На правах рукописи

Коржук Виктория Михайловна

МОДЕЛЬ И МЕТОД ИДЕНТИФИКАЦИИ АТАК СЕТЕВОГО УРОВНЯ НА БЕСПРОВОДНЫЕ СЕНСОРНЫЕ СЕТИ НА ОСНОВЕ ПОВЕДЕНЧЕСКОГО АНАЛИЗА

Специальность 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург – 2019

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский университет ИТМО».

Научный руководитель: кандидат технических наук, доцент
Заколдаев Данил Анатольевич
Университет ИТМО,
декан факультета безопасности информационных технологий

Официальные оппоненты: **Суханов Андрей Вячеславович**
доктор технических наук, профессор
АО «Эврика», советник генерального директора

Красов Андрей Владимирович,
кандидат технических наук, доцент,
ФГБОУ ВО Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, заведующий кафедрой Защищенных систем связи

Ведущая организация: **Федеральное государственное бюджетное образовательное учреждение высшего образования «Государственный университет морского и речного флота имени адмирала С.О. Макарова»**

Защита состоится ___ декабря 2019 г. в ___ часов на заседании диссертационного совета по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук Д 002.199.01, созданного на базе Федерального государственного бюджетного учреждения науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН) по адресу: 199178, Санкт-Петербург, 14-я линия В.О., 39.
Факс: (812)-328-44-50, тел.: (812)-328-34-11.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки Санкт-Петербургского института информатики и автоматизации Российской академии наук по адресу: 199178, Санкт-Петербург, В.О., 14 линия, д. 39 и на сайте http://www.spiiras.nw.ru/dissovet/

Автореферат разослан «___» _____ 2019 года.

Ученый секретарь совета
диссертационного совета Д 002.199.01,
кандидат технических наук

Зайцева
Александра
Алексеевна

Общая характеристика работы

Актуальность темы. Современный этап развития информационных технологий характеризуется повсеместным внедрением и использованием различных киберфизических систем (КФС). В качестве основы для таких систем нередко используются беспроводные сенсорные сети (БСС), и защита информации в этих сетях является новой и актуальной задачей. Важность решения задачи идентификации атак на БСС обусловлена спецификой БСС в КФС и непрерывным ростом количества разнообразных сетевых угроз, реализация которых может привести к финансовым, репутационным и даже человеческим потерям. Согласно долгосрочному прогнозу научно-технологического развития РФ до 2030 года к 2020 году в мире будет 30,1 млрд беспроводных устройств, включая системы автоматизации производства (Индустрия 4.0), системы автоматизации зданий и помещений (Умный дом), носимые устройства, в том числе медицинского назначения, системы мониторинга состояния окружающей среды и т.д.

Текущий уровень развития научно-методического аппарата (НМА) не позволяет обеспечивать необходимый уровень целостности и доступности информации, циркулирующей в БСС. Ввиду таких особенностей функционирования устройств БСС, как малый объем памяти, вычислительных мощностей, а также ограничений в электропитании, существующие методы защиты информации и классические системы обнаружения вторжений (СОВ) оказываются недостаточно эффективными.

Мониторинг состояния сети и обнаружение сетевых атак на основе анализа аномалий является одним из самых популярных методов, однако чаще всего исследователи используют семантическую составляющую передаваемых пакетов в трафике. В таких случаях достаточно велика вероятность ложного срабатывания, так как не представляется возможным точно определить причину аномалии: вызвана ли она атакой, программным сбоем или действительным изменением окружающей среды. Сигнатурный анализ также используется часто, однако здесь возникает вопрос хранения и дополнения базы данных сигнатур. Алгоритмы обнаружения атак, основанные на управлении репутацией или доверии показывают высокую точность, но являются вычислительно сложными, что не подходит под ограничения БСС, связанных с малым объемом памяти, энергоемкости и вычислительных ресурсов. Поэтому в данном диссертационном исследовании используется поведенческий анализ. Под поведением понимается набор характеристик сети в конкретный момент времени. Такой подход является не в полной мере исследованным. В соответствии с этим задача идентификации атак сетевого уровня на беспроводные сенсорные сети является актуальной, а предлагаемый в настоящем диссертационном исследовании модельно-методический аппарат направлен на ее решение.

Степень разработанности темы. Значительный вклад в решение проблемы идентификации атак на компьютерные сети внесли такие отечественные исследователи, как И.А. Зикратов, П.Д. Зегжда, Е.С. Басан, Т.И. Гришечкина, А.А. Браницкий, И.И. Викснин и зарубежные исследователи Ю. Эль Мураби, Г. Калнур, Ш. Зонг, Х. Ку, Ма и др. В частности, задачей обеспечения информационной

безопасности БСС посвящены работы И.В. Котенко, И.Б. Саенко, И.С. Лебедева, Е.Е. Бессоновой и А. Да Силвы, М.Р. Ахмеда, И. Альмомани, С. Сингха и др. Анализ работ в этой области показал, что существующие решения позволяют идентифицировать всего лишь от одной до четырех атак сетевого уровня на БСС, ограничиваются использованием единственного классификатора и не имеют гибкой методики идентификации сетевых атак. Соответственно, диссертационная работа направлена на повышение эффективности идентификации атак на БСС с помощью анализа поведения сети, что и определяет актуальность исследования, его теоретическую и практическую значимость.

Научная задача состоит в разработке и обосновании научно-методического аппарата по идентификации атак сетевого уровня на БСС на основе анализа новой комбинации признаков, характеризующего поведение такой сети.

Объектом исследования являются системы обеспечения информационной безопасности БСС.

Предметом исследования являются модели и методы идентификации атак сетевого уровня на БСС.

Цель диссертационного исследования: повышение эффективности идентификации атак сетевого уровня на БСС при помощи оригинального научно-методического аппарата, основанного на анализе поведения сети. Для достижения цели сформулированы и решены следующие частные задачи:

1) Провести исследование и анализ существующих методов идентификации атак на БСС, выявить показатели эффективности их идентификации.

2) Разработать модель профиля поведения устройств БСС.

3) Разработать метод идентификации атак сетевого уровня на БСС на основе анализа поведения сети с использованием совокупности методов машинного обучения.

4) Разработать методику идентификации атак на БСС, содержащую дополнительные этапы настройки и классификации атак.

5) Провести вычислительный эксперимент и обосновать применимость разработанной модели профиля поведения БСС, метода и методики идентификации атак.

6) Сравнить разработанные модель, метод и методику с существующими исследованиями и сформулировать выводы о результатах диссертационной работы.

Научная новизна диссертации определяется следующим:

1) Модель профиля поведения БСС отличается от известных использованием новой комбинации признаков поведения сети, таких как общее количество пакетов, переданных в сети, максимальное количество отправленных узлом пакетов, максимальное и минимальное количество полученных узлом пакетов и соотношение между количеством созданных и полученных пакетов.

2) Метод идентификации атак сетевого уровня на БСС на основе поведенческого анализа сети отличается от известных применением совокупности алгоритма «случайного леса» и вероятностного классификатора и введением параметра степени уверенности.

3) Методика идентификации атак на беспроводные сенсорные сети отличается от существующих использованием разработанных модели профиля поведения БСС и метода идентификации атак сетевого уровня на БСС, а также использованием программной модели реализации атак, программы оценки информативности и программы вычисления апостериорных распределений дискретного параметра распределения многомерной случайной величины по статистической выборке.

Теоретическая и практическая значимость. Разработанные модель, метод и методика предназначены для повышения эффективности идентификации сетевых атак на БСС, что позволит обеспечить необходимый уровень целостности и доступности информации. Предложенные модель и метод позволяют выявить в среднем на 9 атак больше по сравнению с другими исследованиями. Использование параметра степени уверенности предоставляет возможность принять решение о положительной идентификации атак сетевого уровня на БСС при неполном наборе признаков, что позволяет снизить затраты временных и вычислительных ресурсов. Применение разработанной методики идентификации атак и программных продуктов в системе обнаружения вторжений позволяет повысить гибкость системы и обеспечить необходимый уровень защищенности информационных систем. Результаты диссертационной работы могут быть использованы для дальнейшего развития подходов к обеспечению информационной безопасности (ИБ) в БСС в целом и БСС в контексте КФС в частности.

Методология и методы диссертационного исследования. Для решения сформулированных в работе задач использовались следующие методы исследования: методы теории информационной безопасности и методологии защиты информации, методы теории информации, положения теории вероятности, методы математической статистики и вычислительного эксперимента.

Положения, выносимые на защиту:

1) Разработанная модель профиля поведения беспроводной сенсорной сети позволяет идентифицировать большее, по сравнению с существующими исследованиями, количество атак.

2) Разработанный метод идентификации атак сетевого уровня на БСС на основе анализа профиля поведения позволяет обеспечить высокую точность идентификации при использовании неполного набора признаков профиля поведения.

3) Предложенная методика идентификации атак на беспроводные сенсорные сети позволяет повысить эффективность идентификации атак.

Обоснованность и достоверность полученных результатов достигается использованием апробированного математического аппарата и подтверждается проведением сравнительного анализа с существующими методами; серией практических экспериментов по идентификации атак на БСС; согласованностью результатов, полученных при теоретическом исследовании с результатами проведенных экспериментов, а также непротиворечивостью достигнутых результатов и результатов работ других авторов; практической апробацией результатов в научно-исследовательских проектах, деятельности производственных организаций и одобрением на научно-технических конференциях.

Реализация результатов работы. Результаты, представленные в диссертационной работе, были реализованы в рамках выполнения следующих научно-исследовательских работ: проекта Минобрнауки России «Разработка методов агрегации, нормализации, анализа и визуализации больших массивов гетерогенных структурированных, полуструктурированных и неструктурированных данных для мониторинга и управления безопасностью распределенной сети электронных потребительских устройств №14-604-21-0147 (2014-2016); НИР-ПРИКЛ «Исследование уязвимостей систем для защиты от атак по сторонним каналам» №713552 (2014-2016); грант для студентов вузов, расположенных на территории Санкт-Петербурга, аспирантов вузов, отраслевых и академических институтов, расположенных на территории Санкт-Петербурга (2015); гранта РНФ «Управление инцидентами и противодействие целевым киберфизическим атакам в распределенных крупномасштабных критически важных системах с учетом облачных сервисов и сетей Интернета вещей» №15-11-30029 (2015-2017); НИР «Технологии киберфизических систем: управление, вычисления, безопасность» №617026 (2018); НИР-ФУНД «Исследование перспективных методов и технологий защиты киберпространства в банковской сфере» №59880 (2019); НИР-ФУНД «Методы, модели, методики, алгоритмы, протоколы и приложения для обеспечения информационной безопасности киберфизических систем» №717075 (2017-2019). Результаты работы использовались при разработке системы мониторинга информационной безопасности компании АК «Реактор». Полученные результаты используются в образовательном процессе факультета БИТ Университета ИТМО по направлениям подготовки бакалавриата 10.03.01 и магистратуры 10.04.01 по дисциплинам «Основы информационной безопасности», «Теория информационной безопасности и методология защиты информации», «Комплексные системы защиты информации».

Апробация результатов работы. Основные результаты работы были представлены и обсуждены на следующих конференциях и семинарах:

1) 15th International Conference NEW2AN 2015 and 8th Conference ruSMART;

2) 18th Conference of Open Innovations Association FRUCT (2016);

3) Конференция Information Security and Protection of Information Technology в рамках Conference of Open Innovations Association FRUCT (2015);

4) Молодежная научная школа «Безопасные информационные технологии» в рамках XIV Санкт-Петербургской межрегиональной конференции «Региональная информатика» (2014);

5) IX и X Санкт-Петербургская межрегиональная конференция «Информационная безопасность регионов России» (2015, 2017);

6) IV, V, VI, VII, VIII Конгресс молодых ученых (2014, 2015, 2016, 2017, 2018);

7) XLIV, XLV Научная и учебно-методическая конференция Университета ИТМО (2015, 2016);

8) Круглый стол для победителей конкурсов грантов (Университет ИТМО – по техническому направлению) в рамках Конкурса грантов для студентов вузов, расположенных на территории Санкт-Петербурга, аспирантов вузов, отраслевых

и академических институтов, расположенных на территории Санкт-Петербурга (2015).

Публикации. Основные результаты, полученные в ходе диссертационного исследования, изложены в 17 печатных работах, три из которых опубликованы автором в журналах, рекомендованных ВАК, восемь – в зарубежных изданиях, индексированных в Web of Science и Scopus, шесть – в прочих изданиях. Получено три свидетельства о государственной регистрации программ для ЭВМ.

Личный вклад. Все результаты, представленные в диссертационной работе, получены лично автором в процессе выполнения научно-исследовательской работы.

Структура и объем диссертационной работы. Диссертация состоит из введения, трех глав, заключения и 5 приложений. Основной материал изложен на 184 страницах. Полный объем диссертации составляет 206 страниц с 43 рисунками и 32 таблицами. Список литературы содержит 221 наименование.

Содержание работы

Во введении обоснована актуальность темы диссертационной работы; определены и сформулированы цель и задачи исследования; раскрыты принципы используемых подходов; выявлена научная новизна и практическая ценность полученных результатов; сформулированы положения, выносимые на защиту; приведены сведения об апробации результатов исследования.

Первая глава посвящена анализу современного состояния проблемы обеспечения ИБ в БСС, рассмотрены различные варианты мониторинга состояния сети и обнаружения вторжений в такую сеть. Определены ограничения и допущения работы: исследуются топологии «ячеистая сеть» и «кластерное дерево», используется пассивный вид мониторинга. Приведен обзор подходов и выполнен анализ существующих моделей и методов идентификации атак сетевого уровня на БСС. Определены показатели эффективности идентификации в рамках решаемой задачи.

Научная задача исследования состоит в разработке и обосновании научно-методического аппарата по идентификации атак сетевого уровня на беспроводные сенсорные сети на основе анализа новой комбинации признаков, характеризующего поведение такой сети.

Существует некое множество всевозможных объектов – $X = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ и его подмножество, содержащее исследуемые объекты X_c, множество классов $Y = \{C_0, \dots, C_m\}$, каждый из которых представляет собой набор объектов, являющийся подмножеством O_s, и объединённых некоторой связью: $\forall \mathrm{C}_i \in Y$ ($C_i \subset X_\mathrm{c}$). Также существует некий алгоритм, относящийся к множеству всех алгоритмов A, такой что $a : X_\mathrm{c} \to \{1, \dots, m\}$. Введём вспомогательное отображение $\lambda : C \to \{0, \dots, m\}$ – которая ставит каждому классу его номер (метку): $\forall C_i \in Y\ \lambda(\mathrm{C_i}) = i, i = \overline{0, m}$. Будем считать, что каждому объекту x соответствует один класс, то есть существует неизвестная целевая зависимость $y^* : X_\mathrm{c} \to \lambda(\mathrm{Y})$. В связи с этим множество X_c является дизъюнктивным объединением множеств, объединяющих различные классы $C_i (i = 0, \dots, m)$: $X_\mathrm{c} = \bigvee_{i=0}^{m} C_i$. .

Каждый объект $\boldsymbol{x} \in X_c$ представлен вектором признаков $x = (x_1, \ldots, x_k)$. Дано множество классов поведений сети C: нормальное поведение, поведение под атакой; К – множество признаков поведения сети, и множество исследуемых поведений сети X_c. Выделим из множества К подмножество наиболее информативных признаков: $\mathrm{K}_I \subseteq \mathrm{K}$. Каждое поведение представлено вектором признаков размерностью $k = \#\mathrm{K}_\mathrm{I}$. Пусть тогда $X_{c\mathrm{K}_I}$ – исследуемые объекты, представленные вектором из признаков $k \in K_I$.

С помощью некоего правила R построена обучающая выборка: набор маркированных векторов признаков поведения сети $\Xi_{X_c^L} = \{(\boldsymbol{x'}_i, \overline{c}_i)\}_{i=0}^{M} = \bigcup_{C_i \in C} \bigcup_{\boldsymbol{x} \in C_i \cap X_c^L} (\boldsymbol{x},\ \lambda\,(\mathrm{C}))$, где $X_c^L \subseteq X_c$ – множество обучающих векторов, а $x_i' \in C_{i\overline{c}_i}$ – отношение принадлежности. Для решения задачи идентификации атак на БСС, на базе обучающей выборки $\Xi_{X_c^L}$ нужно решить следующую оптимизационную задачу. Введём дополнительно множество $G = \{K_{I_1}, \ldots, K_{I_k}\}$, где $K_{I_i} \subseteq \mathrm{K}_\mathrm{I}, \#K_{I_i} = i$, и вспомогательную функцию $\psi: G \to \{X_{c_{K_{I_1}}}, \ldots, X_{c_{K_{I_k}}}\}$, где $X_{c_{K_{I_i}}}$ – исследуемые объекты $\boldsymbol{x} \in X_c$, представленные с помощью вектора длиной i. Необходимо при заданном уровне эффективности σ идентификации свести к минимуму количество используемых признаков $k \in \mathrm{K}_I$:

$$\Lambda(a, X_c, \sigma) = \min_{i \in \{1,\ldots,k\}} \{\#K_{I_i} \mid K_{I_i} \in G, \Omega(a, \psi(K_i)) \geq \sigma\},$$

где $\Lambda(a, X_c, \sigma)$ – функция оценки минимального набора признаков поведения, σ – наперёд заданное положительное число.

Сформулирована цель исследования, состоящая в повышении эффективности идентификации атак сетевого уровня на БСС на основе нового набора признаков, характеризующих поведение такой сети, и методов машинного обучения.

Во второй главе представлены модель профиля поведения БСС на основе новой комбинации признаков и метод идентификации атак сетевого уровня на БСС на основе поведенческого анализа.

Разработанная модель профиля поведения сети представлена следующим образом:

$$M=<x_1, x_2, \ldots x_n>, \quad (2)$$

где $x_1, x_2, \ldots x_n$ – признаки сети, характеризующие ее поведение в момент времени. На основании стандарта 802.15.4, протокола ZigBee, существующих исследований и наборов данных, относящихся к БСС, было разработано множество мощностью 51. Проанализированы и охарактеризованы атаки на БСС, выявлено 14 актуальных атак сетевого уровня: «Отказ во сне», «Затопление», «Повторная передача» с вариациями, «Выборочная пересылка» с вариациями, «Воронка», «Подмена» с вариациями, «Атака Сибиллы» и «Червоточина».

Для генерации и сбора статистических данных использовалась разработанная в ходе исследования программная модель проведения атак на БСС на основе среды моделирования OMNeT++. Данная программная модель имеет теоретическое (математические расчеты) и экспериментальное обоснование, апробирована и зарегистрирована. Модель учитывает особенности беспроводного канала

связи, сетевого уровня БСС, способна к различным вариантам адресации и топологии и имитирует атакующие узлы.

Было произведено сокращение признакового пространства на основе оценки информативности с помощью формулы Шеннона, формулы Кульбака и метода накопленных частот, оценки параметра дискретизации и коэффициента корреляции Пирсона. Используемая формула Шеннона имеет следующий вид:

$$I_{Sh}(x) = 1 + \sum_{i=1}^{G}\left(P_i * \sum_{k=1}^{K} P_{i,k} * \log_K P_{i,k}\right), \quad (3)$$

$$P_i = \frac{\sum_{i=1}^{K} m_{i,k}}{N}, \quad P_{i,k} = \frac{m_{i,k}}{\sum_{i=1}^{K} m_{i,k}},$$

где G – количество градаций признака x; K – количество классов; N – количество объектов всех классов; $m_{i,k}$ – количество объектов класса k, у которых признак принимает значение градации i; P_i – частота появления градации i среди всех объектов выборки; $P_{i,k}$ – доля объектов класса k среди всех объектов, у которых признак принимает значение градации i.

Используемая формула Кульбака имеет следующий вид:

$$I_K(x) = \sum_{i=1}^{G}[P_{i1} - P_{i2}] * \log_2 \frac{P_{i1}}{P_{i2}}, \quad (4)$$

$$P_{ik} = \frac{m_{ik}}{\sum_{i=1}^{G} m_{ik}}, k = 1; 2.$$

В результате признаковое пространство было сокращено до 11 признаков. Для автоматизации процесса оценки информативности разработана и зарегистрирована программа для ЭВМ.

Для разработки метода идентификации в качестве основы использовался алгоритм «случайный лес», представляющий собой коллекцию случайных деревьев, показавший наиболее высокий показатель точности (precision), аккуратности (accuracy) и полноты (recall), что не противоречит существующим исследованиям в этой области:

$$F = \{h(X, \Psi_s), s = 1, \ldots\}, \quad (5)$$

где $h(X, \Psi_s)$ – решающее дерево, $\{\Psi_s\}$ – независимые одинаково распределенные случайные векторы, X – выборка признаков x, подающаяся на вход.

В результате обучения алгоритма были выбраны следующие признаки профиля поведения, необходимые для идентификации:

1) num_packets_avg – общее количество пакетов, переданных в сети по спецификации ZigBee, усредненное по числу PAN в сети;

2) num_packets_out_max – максимальное количество пакетов, отправленных каким-либо узлом (собственных и пересланных);

3) num_packets_equal_src_pan_max – максимальное количество полученных каким-либо узлом пакетов, в которых в качестве PAN-отправителя указана одна и та же PAN;

4) num_packets_equal_src_pan_min – минимальное количество полученных каким-либо узлом пакетов, в которых в качестве PAN-отправителя указан один и тот же PAN;

5) frac_packets_created_acquired_avg – соотношение между количеством пакетов, созданных узлом, и количеством полученных пакетов, в которых в качестве источника указан данный узел, усредненное по сети.

За исключением *num_packets_out_max*, указанные признаки являются характерными и информативными только для БСС, прямых аналогов в беспроводных сетях не имеют.

Таким образом, размерность признакового пространства, используемого для построения модели профиля поведения БСС, была снижена до 5 признаков:

$$a\big(L, Inf(L)\big) = L', \tag{6}$$

где a – алгоритм снижения размерности признакового пространства, L – признаковое пространство исходной размерности, *Inf(L)* – суперпозиция функций оценки информативности, дополнительно включающая в себя оценку, производимую «случайным лесом», L' – сокращенное признаковое пространство.

Выдвинуто предположение о возможности идентификации атак на неполном наборе признаков с помощью применения вероятностного классификатора. В контексте классификации больших объемов данных на несколько классов положено, что: x – единичное измерение, которое может быть случайным вектором $\mathbf{x}$, $\mathbf{X} = \{x_1, \dots, x_n\}$ – набор данных, полученный из n независимых одинаково распределенных случайных величин x_i, θ – параметр распределения величины x: $p(x|\theta)$, или класс.

Тогда при предположении, что распределение x меняется в зависимости от класса, вероятность принадлежности измерения классу вычисляется как

$$p(\theta|\mathbf{X}) = \frac{p(\mathbf{X}|\theta)p(\theta)}{\sum_{i=1}^{m} p(\mathbf{X}|\theta_i)p(\theta_i)}, \tag{7}$$

где $P(\theta|\boldsymbol{X})$ – апостериорная вероятность принадлежности измерения классу θ при известных данных $\mathbf{X}$, $p(\boldsymbol{X}|\theta)$ – вероятность получения данных $\mathbf{X}$ в классе θ, $p(\theta)$ – априорная вероятность класса θ, $p(\mathbf{X})$ – вероятность получения данных $\mathbf{X}$. При этом используются нормализующие константы, вытекающие из формулы полной вероятности:

$$\frac{\sum_{i=1}^{m} p(\mathbf{X}|\theta_i)p(\theta_i)}{p(\mathbf{X})} = 1 \ , \tag{8}$$

$$p(\mathbf{X}) = \sum_{i=1}^{m} p(\mathbf{X}|\theta_i)p(\theta_i).$$

Изначально алгоритмом предусматривается расчет апостериорных распределений сочетаний (комбинаций) признаков из неполного набора (подвектора, т.е. определенного признакового пространства классификации) по k элементов. При этом количество таких сочетаний будет равно

$$C = \sum_{i=1}^{t} C_n^{k_i} = \sum_{i=1}^{t} \frac{n!}{k_i!(n-k_i)!} \tag{9}$$

Выражение (9) определяет количество итераций в цикле.

Однако значение C в общем случае зависит экспоненциально от размерности u неполного набора признаков, поэтому

$$f(u) = O(c^u), c \in \left[\sqrt[u]{u}; 2\right]. \quad (10)$$

В лучшем случае $f(u) = O\left(\sqrt[u]{u}^u\right) = O(u)$, если для классификации используется лишь один признак, в худшем – $f(u) = O(2^u)$, если используются все, так как $C = \sum_{k=1}^{n} C_n^k = 2^n - 1$.

Также в процесс идентификации введен параметр степени уверенности *c*, который здесь является субъективным параметром и показывает, насколько можно быть уверенным, что класс наблюдения именно таков, какой был определен. В прикладном смысле применения этот параметр отражает заданную ожидаемую апостериорную вероятность определенного класса: $c = p(\Theta|x)$. Параметр степени уверенности гарантирует среднюю точность классификации, равную ему или превышающую его.

В обобщенном виде процесс идентификации атак на основе вероятностного классификатора представлен на рисунке 1.

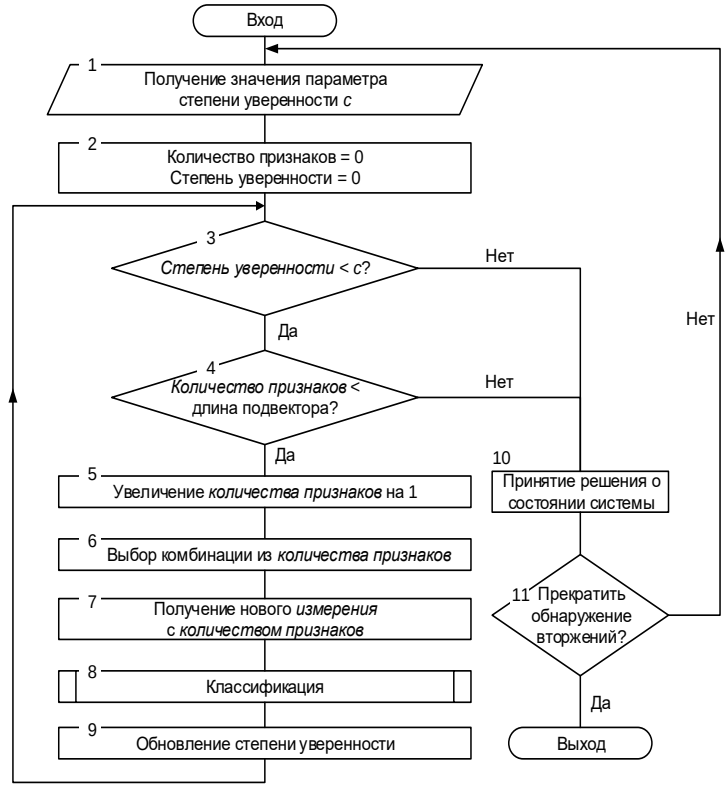

Рисунок 1 – Блок-схема алгоритма обнаружения вторжений на основе вероятностного классификатора

В противном случае, на следующей итерации анализируются два признака. По достижению соответствия параметру степени уверенности идентификация прекращается. Если заданная степень уверенности не достигнута и используются все 5 признаков, ожидается классификация с помощью алгоритма «случайный лес» на периоде *T_b*.

Анализ неполного набора признаков из профиля поведения сети вероятностным классификатором производится каждые *T_s*. В этом случае вероятностный классификатор будет работать с растущим объемом набора данных: *T_s*, *2*T_s*, *3*T_s .. nT_s*, где *nT_s* < *T_b*.

Особенность разработанного метода заключается в последовательном применении вероятностного классификатора и «случайного леса». Вероятностный классификатор используется для анализа неполного набора признаков профиля поведения с учетом заданного параметра степени уверенности и необходим для примерной оценки поведения БСС. Процесс начинается с одного признака. Идентификация считается успешной, если точность идентификации больше или равна показателю степени уверенности.

При невозможности идентификации выводится сообщение о неизвестном аномальном поведении. В таблице 1 приведено количество классификаций по типам поведения.

Таблица 1 – Количество классификаций в процентном выражении

		Индекс определенного поведения														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Индекс поведения	1	99%	-	-	-	-	-	-	-	-	-	-	1%	-	-	-
	2	-	100%	-	-	-	-	-	-	-	-	-	-	-	-	-
	3	1%	-	90%	-	1%	-	-	7%	-	-	-	-	-	-	1%
	4	-	-	-	99%	-	-	-	-	-	-	-	-	-	-	1%
	5	-	-	-	1%	99%	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	99%	-	-	-	-	1%	-	-	-	1%
	7	1%	-	-	-	-	-	99%	-	-	-	-	-	-	-	-
	8	-	-	15%	-	-	-	-	85%	-	-	-	-	-	-	-
	9	1%	-	-	-	-	-	-	1%	98%	-	-	-	-	-	-
	10	-	1%	-	-	-	-	-	-	-	99%	-	-	-	-	-
	11	-	-	-	-	-	-	-	-	-	-	99%	1%	-	-	-
	12	1%	-	-	-	-	-	-	-	-	-	-	99%	-	-	-
	13	-	-	-	-	-	-	-	-	-	-	-	-	100%	-	-
	14	-	-	-	-	1%	-	-	-	-	-	-	-	-	99%	-
	15	-	-	-	-	-	-	-	1%	1%	1%	-	-	-	-	97%

Полученные экспериментальные зависимости средней точности классификации от параметра степени уверенности для различных типов сетевых атак на БСС представлены на рисунке 2. Как видно из этого рисунка, зависимость между параметром степени уверенности и точностью классификации не является строго линейной, однако при возрастании первого возрастает и второе. При этом каждый класс обладает минимально возможной точностью классификации и максимально возможной, к которой точность асимптотически приближается при повышении значения параметра степени уверенности. Следует отметить, что в большинстве случаев при необходимости обеспечения в среднем заданной точности классификации значение данного параметра не должно быть обязательно равно требуемой точности, Он может быть меньше, в пределах 30%.

Таким образом, разработаны: модель профиля поведения беспроводной сенсорной сети, позволяющая идентифицировать 14 атак сетевого уровня; программная модель реализации атак на БСС; метод идентификации атак на основе совместного использования алгоритма «случайный лес» и вероятностного классификатора и введения параметра степени уверенности, позволяющий обеспечивать идентификацию атак сетевого уровня с заданной точностью.

Третья глава посвящена разработке методики идентификации атак сетевого уровня на БСС, экспериментам и сравнению результатов. Содержит практические рекомендации по применению методики.

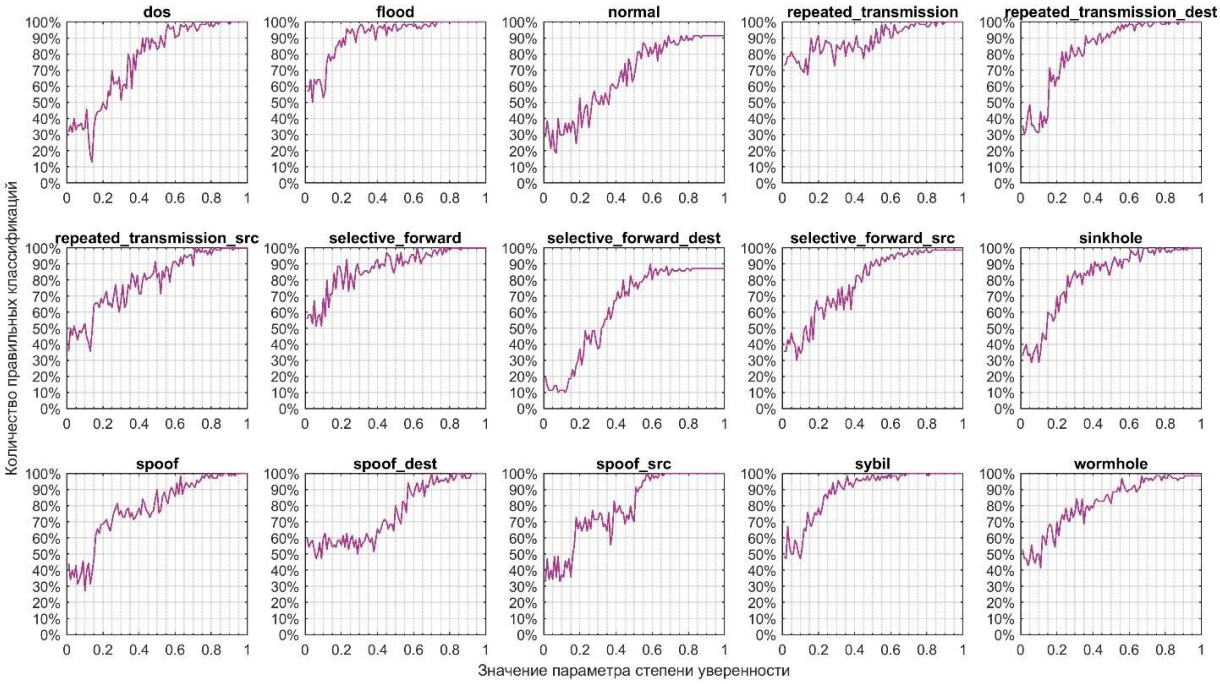

Рисунок 2 – Зависимость средней точности классификации от параметра степени уверенности

Предложенная методика приведена на рисунке 3 и включает в себя этап настройки программной модели атак на БСС, что позволяет задать параметры сети и при необходимости изменить перечень идентифицируемых атак.

Благодаря использованию разработанной модели формируются профили поведения сети. Для оценки информативности и вычисления апостериорных вычислений также используется разработанное ПО. Метод идентификации атак может быть также настроен в соответствии с требованиями и ограничениями сети, что необходимо для повышения уровня целостности и доступности КФС, в основе которых лежат БСС.

Для подтверждения результатов исследования были проведены расширенные эксперименты при изменении топологии и размера выборки, а также при изменении параметра степени уверенности. Итоговая диаграмма по сравнению эффективности идентификации включает в себя результаты расширенных экспериментов. После тестирования метода при фиксированном значении параметра степени уверенности было проведено исследование зависимости от него точности и среднего числа признаков классификации.

Были получены данные о зависимости среднего количества использованных признаков классификации от апостериорной вероятности нормального состояния (рис. 4). Для большинства атак наблюдается относительное постоянство значений зависимости на диапазоне от 0 до 0,7±0,1, после чего происходит резкое снижение. Здесь выделяются состояния repeated_transmission и spoof_dest, для которых графики зависимости демонстрируют практически отсутствие зависимости от априорной вероятности нормального состояния: в их статистике практически (или совсем) отсутствуют значения, походящие на нормальные. Из данных рисунка 4 видно, что зависимость между параметром степени уверенности

и точностью классификации не строго линейная, однако при возрастании первого возрастает и второе.

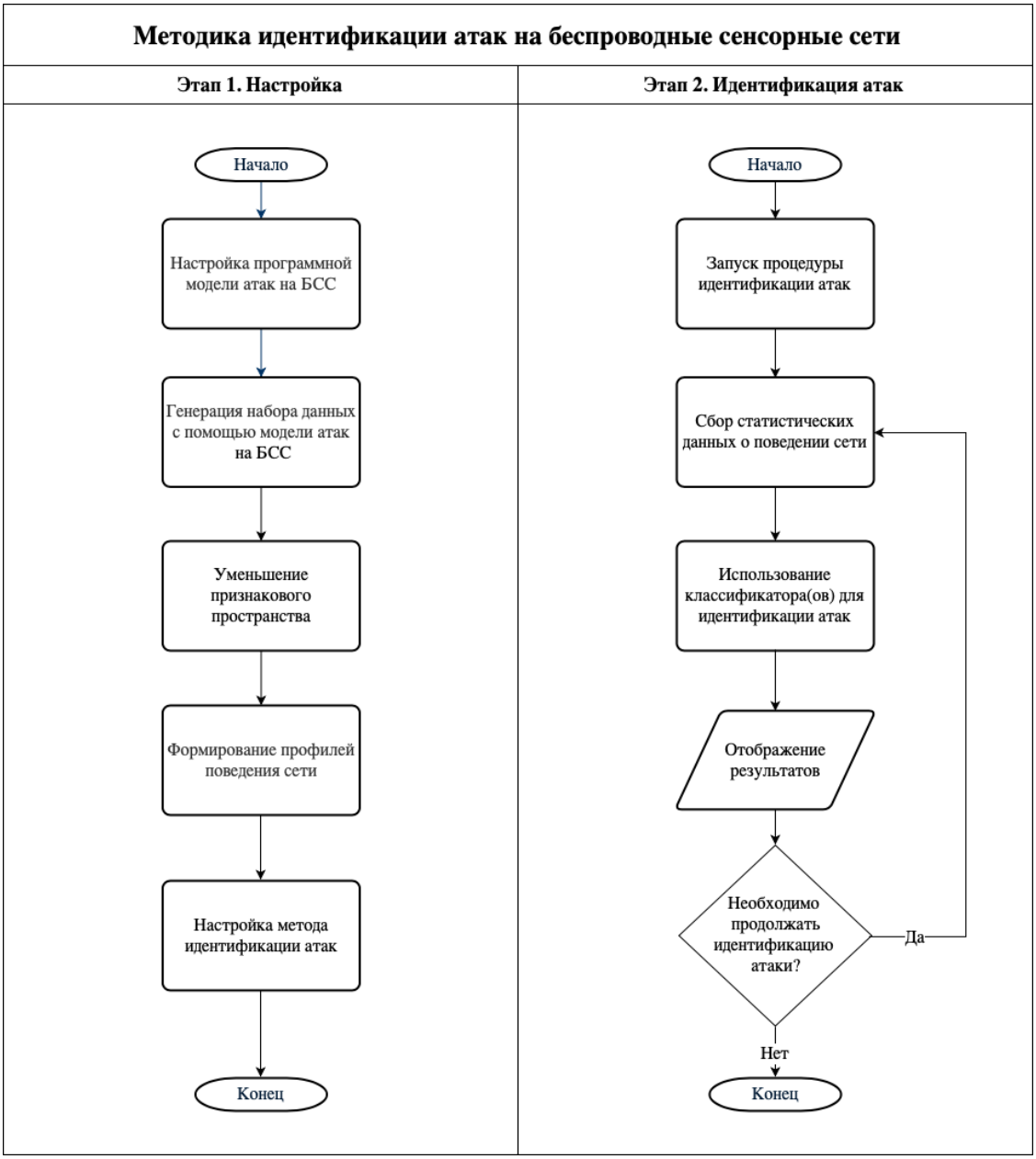

Рисунок 3 – Методика идентификации атак на БСС

Для обеспечения высокой точности идентификации было установлено, что значение параметра степени уверенности должно быть выше значения апостериорной вероятности нормального состояния приблизительно на 30%; нет необходимости в использовании апостериорной вероятности нормального состояния, превышающей 20%. Результаты экспериментов показали, что отсутствует необходимость более чем в 3 признаках классификации для всех атак, кроме selective_forward_dest при любых значениях параметра степени уверенности и априорной вероятности нормального состояния.

На рисунке 5 представлено сравнение эффективности идентификации с существующими исследованиями. Прирост эффективности составил около 20% по 4 показателям.

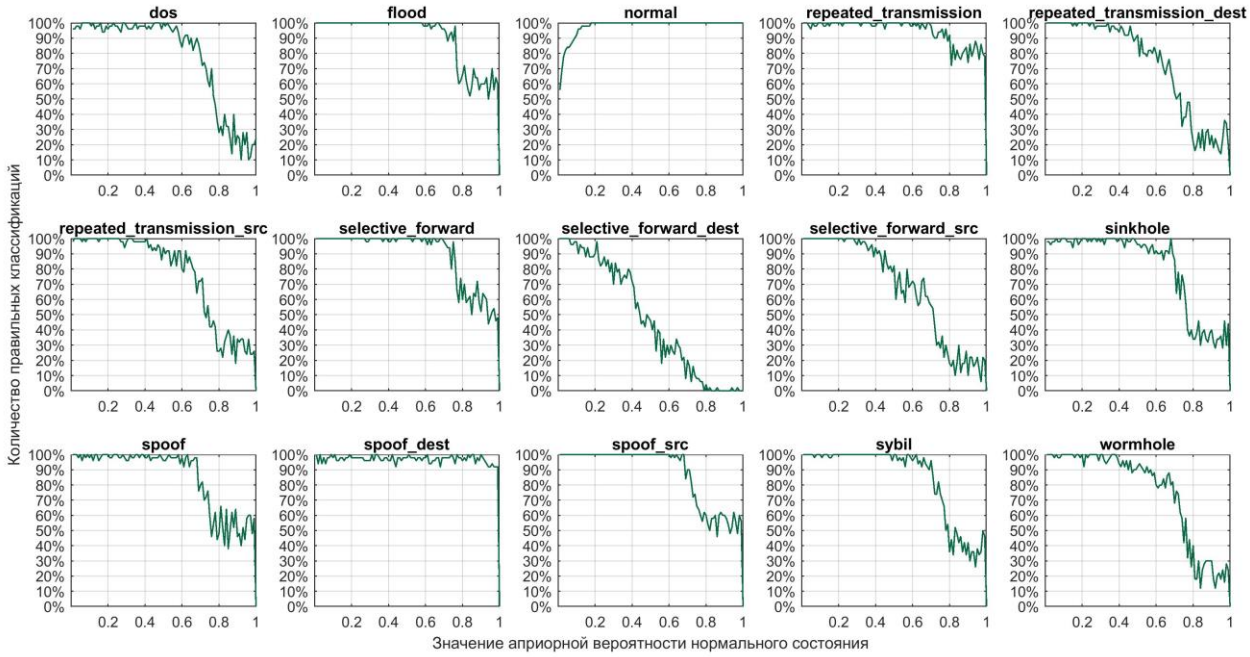

Рисунок 4 – Зависимость количества правильных классификаций от значения априорной вероятности нормального состояния

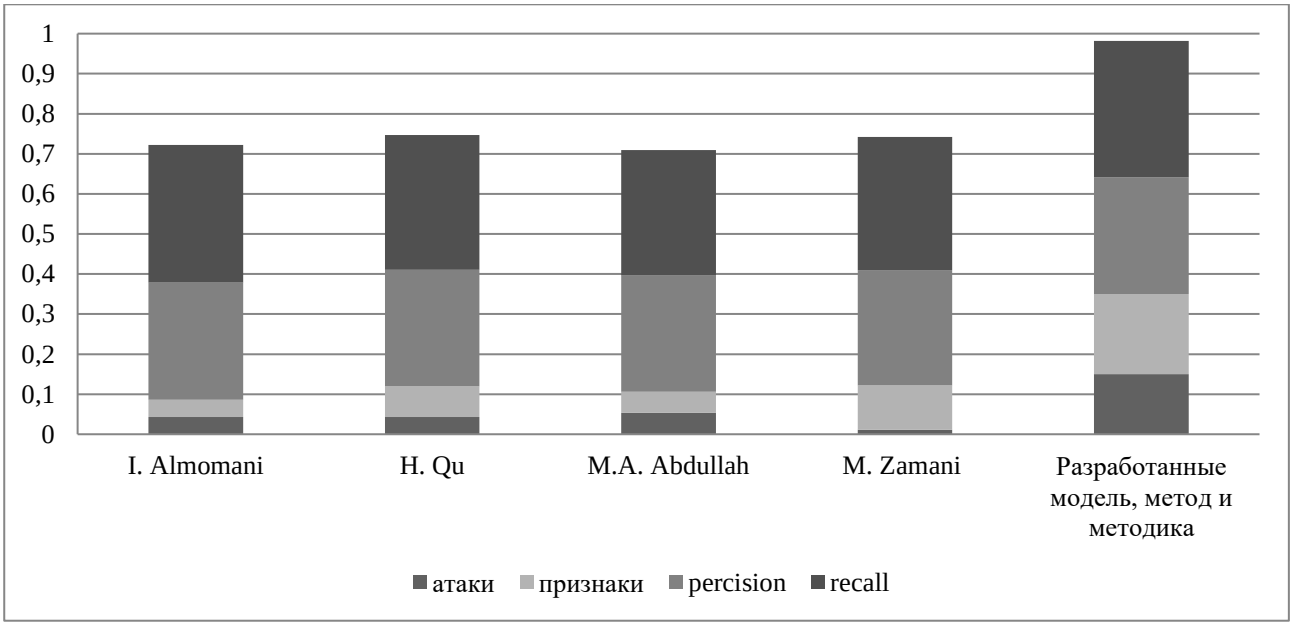

Рисунок 5 – Сравнение эффективности разработанных модели, метода и методики с существующими работами в области идентификации атак на БСС

Сделаны выводы о том, что разработанные модель профиля поведения сети, метод и методика идентификации атак удовлетворяет поставленным в работе задачам и обеспечивают повышение уровня целостности и доступности информационных ресурсов в БСС.

В качестве практической реализации приведена возможность использования разработанных модели, метода и методики как составляющей части системы мониторинга состояния сети и СОВ в КФС, имеющих в своем составе БСС.

В **заключении** приведены основные научные результаты диссертационной работы и описаны перспективы дальнейшего исследования в рамках тематики.

Основные выводы и результаты

В диссертационной работе решена задача разработки научно-методического аппарата для идентификации атак сетевого уровня на БСС на основе поведенческого анализа, имеющая большое значение для обеспечения ИБ КФС. Основные результаты представлены ниже:

1. Рассмотрены существующие методы идентификации атак на БСС, проведен их анализ, выявлены достоинства и недостатки данных методов. Проанализированы условия и ограничения применения каждого из методов.

2. Разработана модель профиля поведения беспроводной сенсорной сети, позволяющая идентифицировать 14 атак сетевого уровня на БСС.

3. Разработана программная модель реализации атак, позволяющая адаптировать разработанную модель к конкретным условиям и получить модельные статистические данные, что упрощает процесс обеспечения информационной безопасности БСС.

4. Разработан метод идентификации атак сетевого уровня на беспроводные сенсорные сети на основе анализа профиля поведения сети, использующий комбинацию алгоритмов «случайный лес» и вероятностного классификатора, обладающий более высокой точностью по сравнению с исследованными методами, использующий неполный набор признаков и позволяющий задавать необходимую точность идентифицировать атаки на БСС.

5. Разработана методика идентификации атак на беспроводные сенсорные сети, позволяющая повысить эффективности идентификации и, соответственно, уровень защищенности БСС благодаря гибкой настройке на основе разработанных модели профиля поведения, программной модели проведения атак и метода идентификации атак, программы подсчета информативности признаков статистической выборки и программы вычисления апостериорных распределений дискретного параметра распределения многомерной случайной величины по статистической выборке.

Представленный научно-методический аппарат для идентификации атак сетевого уровня на беспроводные сенсорные сети на основе анализа поведения такой сети позволяет повысить эффективность идентификации атак на БСС и может быть использован в качестве основы для построения системы обнаружения вторжений.

Рекомендации по применению разработанного научно-методического аппарата для идентификации атак сетевого уровня на БСС включают в себя указания по использованию модели профиля поведения БСС, настройке программной модели реализации атак на БСС и применению метода и методики идентификации атак, а также разработку подходов, направленных на повышение эффективности процесса идентификации атак. Основываясь на сформулированных рекомендациях, представляется вероятной возможность применения полученных результатов в системах обнаружения вторжений в целом и в рамках концепции кибер-физических систем в частности, так как разработанные модель профиля поведения, метод и методика идентификации позволяют обнаружить нарушения целостности и доступности информации, циркулирующей в таких системах.

В качестве **перспектив** дальнейшей разработки тематики следует указать исследования, связанные с развитием модели профиля поведения беспроводной сенсорной сети и доработки программной модели реализации атак на БСС для увеличения количества идентифицируемых атак и улучшения показателей идентификации. Другим направлением является расширение и апробация разработанных модели профиля поведения, программной модели реализации атак и метода идентификации для различных сетевых протоколов в БСС. Кроме того, представляется значимой разработка методов противодействия исследованным атакам. Также возможно использование описанных в работе модели, метода и методики для разработки мобильных программно-аппаратных средств мониторинга состояния БСС.

Полученные результаты соответствуют п. 3 «Методы, модели и средства выявления, идентификации и классификации угроз нарушения информационной безопасности объектов различного вида и класса» и п. 14 «Модели, методы и средства обеспечения внутреннего аудита и мониторинга состояния объекта, находящегося под воздействием угроз нарушения его информационной безопасности» по специальности 05.13.19 паспорта специальностей ВАК (технические науки).

Список публикаций по теме диссертации

В журналах, рекомендованных ВАК

1. Коржук В.М. Обеспечение информационной безопасности каналов связи на основе многофункционального специализированного программно-аппаратного решения / В.М. Коржук, М.Е. Сухопаров, И.С. Лебедев, И.Е. Кривцова, С.А. Печеркин // Проблемы информационной безопасности. Компьютерные системы - 2016. - № 2. - С. 70-74 (№1697 в списке ВАК, ISSN 2071-8217)

2. Коржук В.М. Идентификация атак на беспроводные сенсорные сети на основе анализа аномального поведения сети / В.М. Коржук, П. Бонковски // Научно-технический вестник Поволжья - 2018. - № 2. - С. 83-85 (№1446 в списке ВАК ISSN 2079-5920)

3. Коржук В.М. Введение параметра степени уверенности в процесс идентификации атак на киберфизические системы / В.М. Коржук, А.В. Грозных, Д.А. Заколдаев // Современная наука: актуальные проблемы теории и практики. Серия «Естественные и технические науки» - 2019. - №10. – С. 57-64 (№1926 в списке ВАК, ISSN 2223-2966).

В зарубежных изданиях, индексированных в Web of Science и Scopus

4. Lebedev I.S., Korzhuk V.M. The Monitoring of Information Security of Remote Devices of Wireless Networks // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) - 2015, Vol. 9247, pp. 3-10.

5. Zikratov I.A., Lebedev I.S., Korzhuk V.M. The Estimation of Secure Condition of Multi-Agent Robotic System in Case of Information Influence on the Single Element // Proceedings of the 17th Conference of Open Innovations Association FRUCT - 2015, pp. 362-367.

6. Lebedev I.S., Korzhuk V., Krivtsova I., Salakhutdinova K., Sukhoparov M.E., Tikhonov D. Using Preventive Measures for the Purpose of Assuring Information Security of Wireless Communication Channels // Proceedings of the 18th Conference of Open Innovations Association FRUCT - 2016, pp. 167-173.

7. Lebedev I.S., Krivtsova I.E., Korzhuk V., Bazhayev N., Sukhoparov M.E., Pecherkin S., Salakhutdinova K. The Analysis of Abnormal Behavior of the System Local Segment on the Basis

of Statistical Data Obtained from the Network Infrastructure Monitoring // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) - 2016, Vol. 9870, pp. 503-511.

8. Korzhuk V., Krivtsova I., Shilov I. The Model of the Attack Implementation on Wireless Sensor Networks // Proceedings of the 20th Conference of Open Innovations Association FRUCT - 2017, pp. 187-194.

9. Zikratov I.A., Korzhuk V., Shilov I., Gvozdev A. Formalization of the Feature Space for Detection of Attacks on Wireless Sensor Networks // Proceedings of the 20th Conference of Open Innovations Association FRUCT - 2017, pp. 526-533

10. Korzhuk V., Shilov I., Torshenko J. Reduction of the Feature Space for the Detection of Attacks of Wireles Sensor Networks // Proceedings of the 20th Conference of Open Innovations Association FRUCT - 2017, pp. 195-201.

11. Korzhuk V., Groznykh A., Menshikov A., Strecker M. Identification of Attacks against Wireless Sensor Networks Based on Behaviour Analysis // Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications - 2019, Vol. 10, No. 2, pp. 1-21.

В прочих изданиях

12. Коржук В.М. Об актуальности развития и обеспечения безопасности сети ZigBee // Региональная информатика (РИ-2014). XIV Санкт-Петербургская международная конференция «Региональная информатика (РИ-2014)» – 2014. - С. 568-569.

13. Коржук В.М. Анализ статистических закономерностей откликов от удаленных устройств беспроводных сетей в целях мониторинга информационной безопасности. - 2015. – Т. Сборник тезисов докладов конгресса молодых ученых. Электронное издание. – СПб: Университет ИТМО, 2015.

14. Коржук В.М. Исследуемые методы оценки информационной безопасности сенсорных сетей // Сборник тезисов «XLV научная и учебно-методическая конференция» Университета ИМТО – 2016.

15. Коржук В.М. Защита беспроводной сенсорной сети от атак маршрутизации (Routing Attacks)/ В.М. Коржук, Д.А. Мочалов // Сборник тезисов V конгресса молодых ученых – 2016.

16. Коржук В.М. Использование статистических данных об инфраструктуре сети для анализа аномального поведения локального сегмента системы // Научные работы участников конкурса «Молодые ученые Университета ИТМО» 2016 года. – СПб.: Университет ИТМО, 2017. – 501 с. – 2017.

17. Коржук В.М., Оценка аномального поведения узлов беспроводной сенсорной сети на основе статистических методов / В.М. Коржук, И.М. Шилов // Сборник тезисов докладов конгресса молодых ученых. Электронное издание. – 2017.

Свидетельства о государственной регистрации программ для ЭВМ

18. Коржук В.М. Программная модель атак на беспроводные сенсорные сети ZigBee / В.М. Коржук, А.А. Воробьева, И.М. Шилов. – Свидетельство о государственной регистрации программы для ЭВМ № 2018617190 от 20.06.2018.

19. Коржук В.М. Программа подсчета информативности признаков статистической выборки / В.М. Коржук, А.А. Воробьева, И.М. Шилов. – Свидетельство о государственной регистрации программы для ЭВМ № 2018618975 от 24.07.2018.

20. Коржук В.М. Программа вычисления апостериорных распределений дискретного параметра распределения многомерной случайной величины по статистической выборке / В.М. Коржук, А.А. Воробьева, А.В. Грозных. – Свидетельство о государственной регистрации программы для ЭВМ № 2018619014 от 25.07.2018.