ВВЕДЕНИЕ ПАРАМЕТРА СТЕПЕНИ УВЕРЕННОСТИ В ПРОЦЕСС ИДЕНТИФИКАЦИИ АТАК НА КИБЕРФИЗИЧЕСКИЕ СИСТЕМЫ

USING THE CONFIDENCE LEVEL PARAMETER IN THE PROCESS OF ATTACK IDENTIFICATION IN CYBER-PHYSICAL SYSTEMS

V. Korzhuk
A. Groznykh
D. Zakoldaev

Summary. This article describes the problem of attack identification in cyber-physical systems. A method for identifying of network-level attacks on wireless sensor networks based on behavioral analysis is described. A variation of the method with the confidence level parameter is proposed. This parameter allows to set the acceptable level of accuracy of attack identification and the number of features used. The results of experiments for 14 different attacks and different network topologies are presented. The values of a priori probability and the confidence level parameter are calculated algorithmically. It was found that the excess of the value of the confidence level parameter over the value of the a posteriori probability of the normal state by ~30% is sufficient; there is no need to use the a posteriori probability of the normal state exceeding 20%. Recommendations on the application of the proposed solution are formulated. The obtained results can be used as a part of intrusion detection system.

Keywords: cyber-physical systems, information security, cyber security, information security, wireless sensor networks, ZigBee, attack identification, confidentiality level.

Коржук Виктория Михайловна
Ассистент, ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»
vika@cit.ifmo.ru
Грозных Антон Владимирович
ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»
groznykhanton@yandex.ru
Заколдаев Данил Анатольевич
К.т.н., доцент, ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»
d.zakoldaev@itmo.ru

Аннотация. В данной статье описана проблема идентификации атак в киберфизических системах. Описан метод идентификации атак сетевого уровня на беспроводные сенсорные сети на основе поведенческого анализа. Предложена вариация метода с использованием параметра степени уверенности, позволяющая задавать допустимый уровень точности идентификации атак и количество используемых признаков. Представлены результаты экспериментов для 8 типов атак с 6 вариациями. Алгоритмически рассчитаны значения априорной вероятности и параметра степени уверенности. Выявлено, что достаточным является превышение значения параметра степени уверенности над значением апостериорной вероятности нормального состояния на ~30%; нет необходимости в использовании апостериорной вероятности нормального состояния, превышающей 20%. Сформулированы рекомендации по применению предложенного решения. Полученные результаты могут применяться в составе систем обнаружения вторжений.

Ключевые слова: киберфизические системы, кибер-безопасность, информационная безопасность, беспроводные сенсорные сети, ZigBee, идентификация атак, степень уверенности.

Введение

Проблема обеспечения информационной безопасности киберфизических систем в современном мире становится все более актуальной. Для беспроводных сенсорных сетей, которые часто используются в качестве основы для таких систем, существует ряд ограничений, касающихся объема памяти, ресурсопотребления и энергоемкости [4]. Для своевременного реагирования на зловредное воздействие и, соответственно, обеспечения информационной безопасности в беспроводных сенсорных сетях, необходимо в режиме реального времени проводить мониторинг состояния сети. Существующие исследования в этой области в основном направлены на выявление аномального поведения или однозначную идентификацию атак [7]. Системы обнаружения вторжений, используемые для локальных или приватных вычислительных сетей, из-за упомянутых выше ограничений не могут также эффективно действовать в беспроводных сенсорных сетях.

Предшествующие исследования

По данным Positive Technologies и в соответствии с моделью угроз, составленной на основе базы данных угроз ФСТЭК, наибольший интерес для злоумышленников представляют атаки на сетевой уровень (в соответствии с моделью OSI), приводящие к нарушению доступности

и целостности информации. Среди них выделяются атаки, приводящие к отказу в обслуживании и изменению маршрутной информации. В исследованиях авторов данной статьи рассматриваются следующие атаки:

1. Атака Сибиллы — происходит создание большого количества зловредных узлов, которые выдают себя за легитимные.
2. «Воронка» — «перетягивание» всего трафика на скомпрометированный узел.
3. «Выборочная пересылка» (три вариации) — позволяет отбросить каждый k-й пакет вместо пересылки.
4. «Затопление» — осуществляется передача пакетов на конкретный адрес с частотой выше, чем в нормальном состоянии.
5. «Отказ во сне» — атакующий узел генерирует и отправляет пакеты определенному узлу в сети, что приводит к истощению источника питания узла.
6. «Повторная передача» (три вариации) — для этой атаки характерна пересылка некоторых пакетов: каждый k-й пакет дублируется во внутренней очереди, а повторная передача из очереди каждого следующего пакета происходит каждые L секунд.
7. «Подмена» (три вариации) — представляет собой подмену адресов отправителя и получателя на случайные для каждого k-го генерируемого пакета.
8. «Червоточина» — характеризуется построением «туннеля» и ретрансляцией пакетов злоумышленников в желаемом направлении.

Ранее авторами были частично исследованы проблемы идентификации атак на беспроводные сенсорные сети на основе спецификации ZigBee [2, 5, 6, 8, 9]. В частности, была разработана модель профиля поведения беспроводной сенсорной сети, отличающаяся новой комбинацией признаков и позволяющей идентифицировать в среднем на 5 атак больше, чем в существующих исследованиях. Ниже представлен перечень признаков, формирующих профиль поведения сети:

- num_packets_avg — общее количество пакетов, переданных в сети по спецификации ZigBee, усредненное по числу PAN в сети.
- num_packet_out_max — максимальное количество пакетов, отправленных каким-либо узлом (собственных и пересланных).
- 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.
- 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.
- 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.

Также, был представлен метод идентификации атак сетевого уровня на основе анализа поведенческих характеристик сети. В данной работе идентификация атак и обнаружение вторжений выступают в качестве синонимичных понятий. Указанный метод использует вышеупомянутую модель профиля поведения сети и алгоритм машинного обучения «случайный лес» для классификации поведения беспроводной сенсорной сети и достигает точности 97%.

Особенностью данных исследований является применение теоретически и экспериментально обоснованной программной модели атак на беспроводные сенсорные сети (№ 2018617190 от 20.06.2018) [5], позволяющей осуществить гибкую настройку метода идентификации, и усовершенствованного метода оценки информативности признаков, также имеющего программную реализацию (Программа подсчета информативности признаков статистической выборки № 2018618975 от 24.07.2018)

Параметр степени уверенности

В результате применения алгоритма «случайный лес» и модели профиля поведения беспроводной сенсорной сети было выявлено, что для достижения высокой точности необходимо использовать все пять признаков модели, при этом период сбора статистических данных о поведении сети должен быть не менее 3600 секунд при допущении, что среднее время генерации пакета составляет 10 секунд [2, 6]. Несмотря на относительно неплохие результаты, было принято решение исследовать возможность использования неполного набора признаков из профиля поведения беспроводной сенсорной сети.

Требованиями к улучшаемому методу идентификации явились обеспечение приемлемой точность обнаружения в обозримые временные промежутки и отсутствие необходимости выделения значительных вычислительных ресурсов. Далее в статье описан метод идентификации атак на основе вероятностной классификации с учетом требований переносимости, масштабируемости, возможности изменения параметров, контролирующих работу метода, а также проанализирована зависимость характеристик метода от указанных параметров.

В качестве основного вычислительного блока используется вероятностный классификатор на основе те-

оремы Байеса [1, 3]. Однако, В контексте классификации больших объемов данных на несколько классов формула приобретает иную интерпретацию. Положим, что

x — единичное измерение, которое может быть случайным вектором **x**; $\mathbf{X} = \{x_1, \ldots, x_n\}$ — набор данных, полученный из n независимых одинаково распределенных случайных величин x_i;

θ — параметр распределения величины $x \sim p(x \mid \theta)$, или класс.

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

$$p(\theta \mid \mathbf{X}) = \frac{p(\mathbf{X} \mid \theta) p(\theta)}{p(\mathbf{X})} \quad (1)$$

где $P(\theta \mid \mathbf{X})$ — апостериорная вероятность принадлежности измерения классу θ при известных данных **X**,

$p(\mathbf{X} \mid \theta)$ — вероятность получения данных **X** в классе θ,

$p(\theta)$ — априорная вероятность класса θ,

$p(\mathbf{X})$ — вероятность получения данных **X**.

При классификации на основе теоремы Байеса генерируется достаточно большая обучающая выборка **X** с назначенными классами θ, на основе которой вычисляются значения $p(\mathbf{X} \mid \theta)$, а также определяются априорные вероятности классов. В результате применения (1) ко всей обучающей выборке вычисляется матрица вида

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nm} \end{bmatrix}, \quad (2)$$

где n — количество уникальных значений в выборке,

m — количество классов.

В случае, если классификация проводится по нескольким признакам, т.е.

$$\mathbf{x} = (x_1, \ldots, x_u)$$
$$\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \quad (3)$$

возникают проблемы, связанные с вычислением члена $p(\mathbf{X} \mid \theta)$, который может быть переписан в соответствии с правилом разложения условных вероятностей:

$$p(x_1, \ldots, x_u \mid \theta) =$$
$$= p(x_1 \mid x_2, \ldots, x_u, \theta)\, p(x_2 \mid x_3, \ldots, x_u, \theta) \ldots p(x_u \mid \theta) \quad (4)$$

Вычисление каждого члена в разложении (4) может быть слишком сложным и затратным по времени и памяти. В наивном Байесовском классификаторе для упрощения вычислений делается предположение о независимости, а именно что каждый признак x_i независим от любого другого x_j, $i \neq j$:

$$p(x_i \mid x_{i+1}, \ldots, x_u, \theta) = p(x_i \mid \theta) \quad (5)$$

В этом случае

$$p(\theta \mid x_1, \ldots, x_u) = \frac{p(\theta, x_1, \ldots, x_u)}{p(x_1, \ldots, x_u)} \propto p(\theta, x_1, \ldots, x_u) =$$
$$= p(\theta) \prod_{i=1}^{n} p(x_i \mid \theta) \quad (6)$$

Под параметром степени уверенности здесь понимается субъективная величина, показывающая насколько можно быть уверенным, что класс наблюдения именно таков, какой был определен. В прикладном смысле применения вероятностного классификатора этот параметр отражает апостериорную вероятность определенного класса. Для автоматизации процесса была разработана программа вычисления апостериорных распределений дискретного параметра распределения многомерной случайной величины по статистической выборки (№ 2018619014 от 25.07.2018).

В обобщенном виде улучшенный метод представлен блок-схемой на рисунке 1. В первую очередь метод получает значение параметра степени уверенности (элемент 1), в общем случае — дополнительными средствами системы обнаружения вторжений (СОВ). Инициализация цикла происходит за счет обнуления значений двух ключевых параметров: количества признаков подпространства и показателя степени уверенности (элемент 2). Перед каждой следующей классификацией происходит проверка удовлетворения двух стоп-правил: достижения значения параметра степени уверенности (элемент 3) — в первую очередь, максимального расширения подпространства (элемент 4) — во вторую. Срабатывание стоп-правила приводит к принятию решения о состоянии, в котором на данный момент находится сеть (элемент 10). Это состояние определяется как класс, присвоенный последнему наблюдению, за которым последовало срабатывание стоп-правила.

В случае, если наблюдение проклассифицировано с низкой степенью уверенности и его подпространство можно расширить, количество признаков увеличивается (элемент 5). По нему выбирается комбинация (элемент 6), значения признаков которой запрашиваются методом: такая процедура позволит ограничиться сбо-

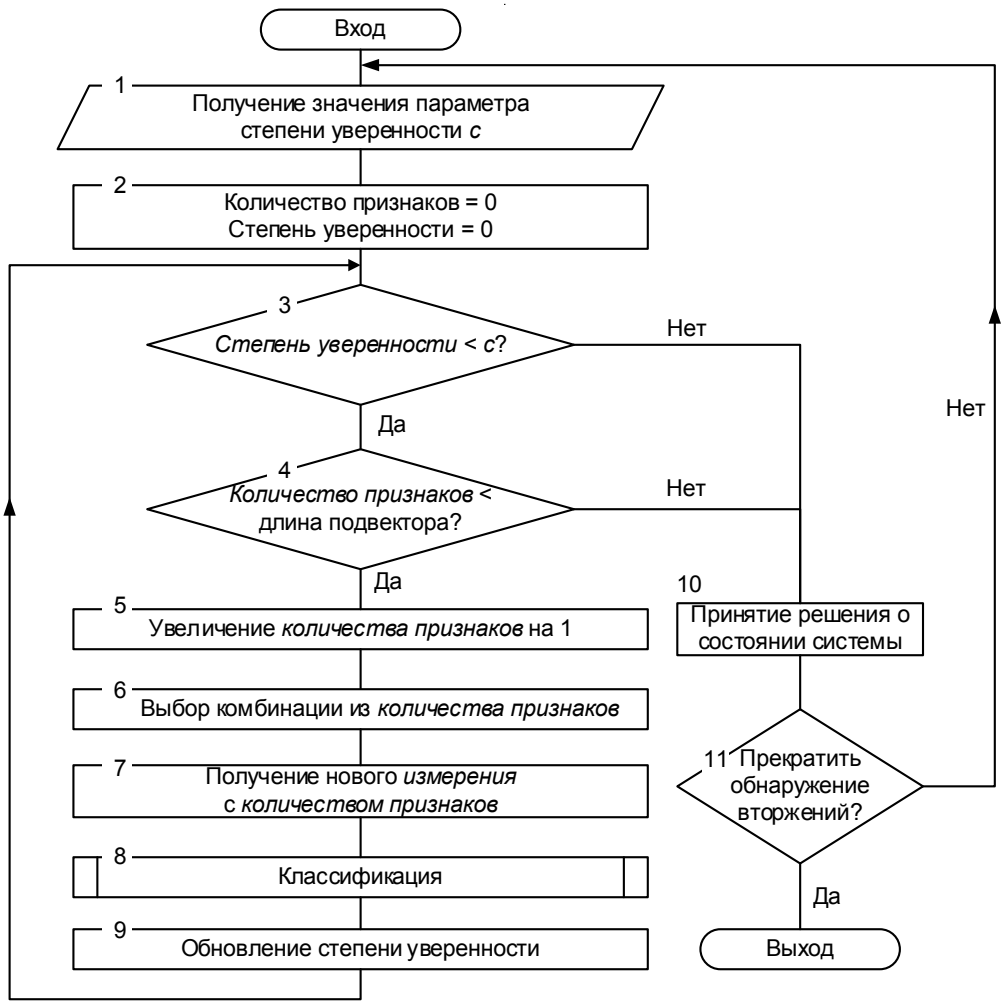

Рис. 1. Блок-схема разработанного метода обнаружения вторжений на основе вероятностного классификатора

ром (элемент 7) сокращенного количества данных с самой БСС, уменьшая нагрузку на ее ресурсы.

Классификация (элемент 8) и обновление степени уверенности производятся на базовой станции БСС, где были собраны данные о работе сети и хранятся таблицы апостериорных вероятностей. Такой метод организации СОВ обладает рядом преимуществ, среди которых:

1. отсутствие высоких требований от вычислительных ресурсов сенсорных узлов;
2. возможность собирать статистику работы БСС на носители данных и анализировать ее;
3. облегченная модификация параметров классификации или вероятностного классификатора.

Главный недостаток такой схемы — централизация метода обнаружения вторжений в одном месте, из-за чего безопасность информации может быть подвергнута угрозе при компрометации базовой станции или нарушению доступности, целостности данных на ней.

Метод проектировался для постоянной работы, т.е. непрерывного мониторинга безопасности БСС на предмет атак, однако после принятия каждого решения (элемент 10) предусмотрена возможность завершить его работу (элемент 11), что выполняется средствами администрирования СОВ. Такая остановка может быть необходима, например, для изменения конфигурации сети, СОВ, вероятностного классификатора или параметра метода идентификации атак.

На рисунке ниже представлены графики изменения апостериорной вероятности определенного класса при увеличении размерности пространства для нескольких серий работы алгоритма итеративной классификации.

Из графиков видна не только общая тенденция повышения степени уверенности, но и ее скорость, а также те состояния, которые требуют большего количества итераций. На данном рисунке также видно, как в состоя-

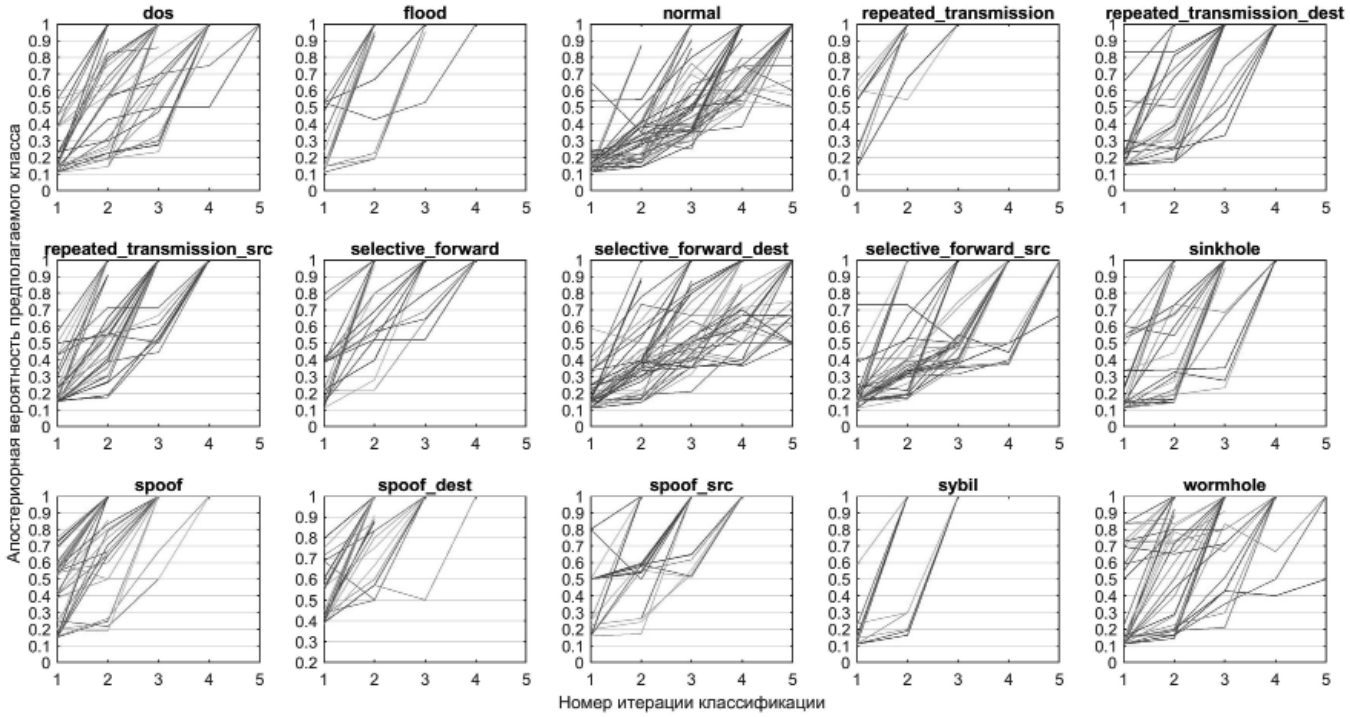

Рис. 2. Графики изменения апостериорной вероятности предполагаемого класса по итерациям для нескольких классификации

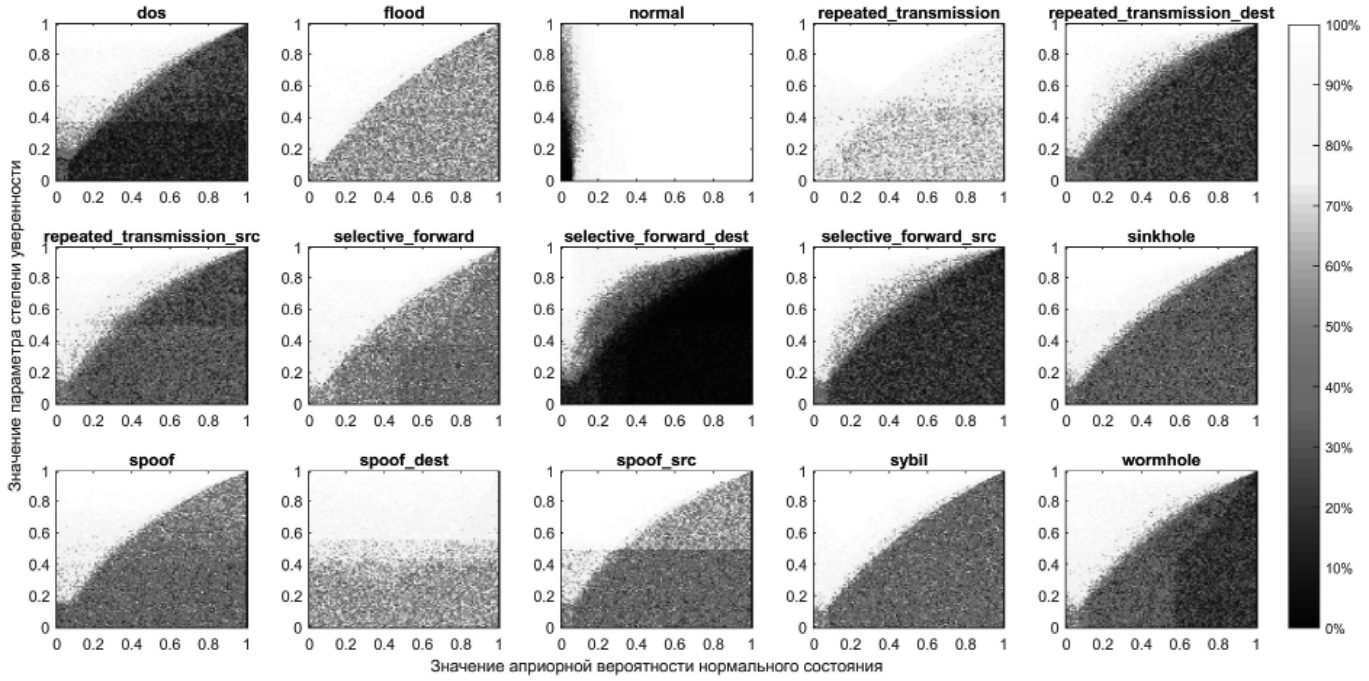

Рис. 3. Графические представления матриц зависимости точности классификации от параметра степени уверенности и априорной вероятности нормального состояния

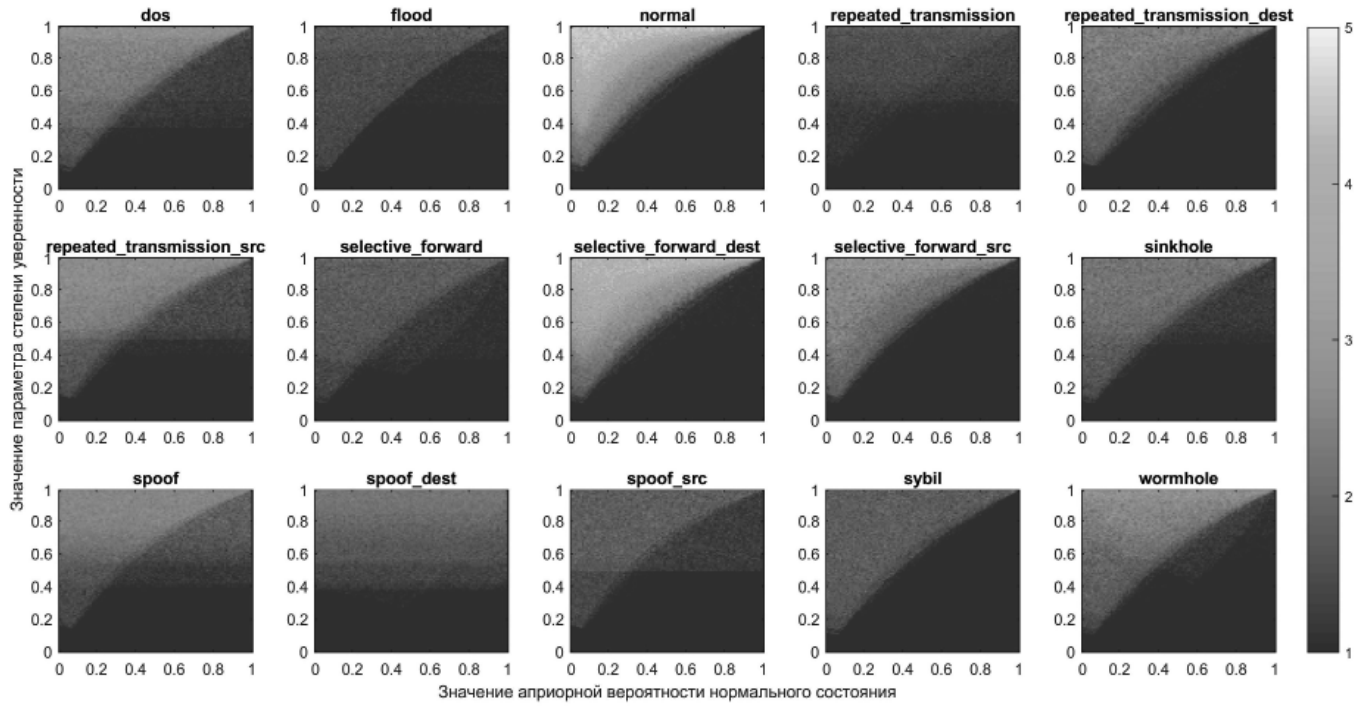

Рис. 4. Графические представления матриц зависимости среднего числа использованных признаков классификации от параметра степени уверенности и априорной вероятности нормального состояния

ниях normal и selective_forward_dest некоторые измерения при использовании пяти признаков классифицируются с более низкой степенью уверенности.

На рисунке 3 изображены графические представления матриц зависимости точности классификации от параметра степени уверенности и априорной вероятности нормального состояния. Каждое значение в матрице определяет цвет, закрашивающий точку, координатами которой являются значения параметров.

Необходимо сказать о тенденции, усматриваемой почти для всех видов атак: для обеспечения высокой точности классификации необходимо, чтобы значение параметра степени уверенности было выше значения апостериорной вероятности нормального состояния на ~30%. Области высокой точности на рисунке — светлые треугольники в левом верхнем углу. Матрица для нормального состояния отличается от других, но только потому, что его апостериорная вероятность была опорной при исследовании. Значения матрицы подтверждают ранее сделанное наблюдение: *нет необходимости в использовании апостериорной вероятности нормального состояния, превышающей 0,2*.

Были также получены аналогичные графические представления матриц для среднего количества признаков классификации (см. рисунок 4. Наиболее важное свойство данных, проиллюстрированное на данном рисунке,— это *отсутствие необходимости более чем в ≈3 признаках классификации для всех атак, кроме selective_forward_dest при любых значениях параметра степени уверенности и априорной вероятности нормального состояния.* Нормальное состояние в некоторых случаях требовало всех пяти признаков.

Необходимо отметить, что для большинства атак наблюдается относительное постоянство значений зависимости среднего количества использованных признаков классификации от апостериорной вероятности нормального состояния на диапазоне от 0 до 70±10%.

Рекомендации по выбору параметров метода идентификации атак

В качестве рекомендаций по применению метода идентификации атак сетевого уровня на БСС с применением параметра степени уверенности были сделаны следующие выводы:

1. для обеспечения высокой точности классификации нормального поведения сети и 14 атак необходимо, чтобы значение параметра степени уверенности было выше значения апостериорной вероятности нормального состояния на ~30%;

2. нет необходимости в использовании апостериорной вероятности нормального состояния, превышающей 20%.

Выводы

Введение параметра степени уверенности в метод идентификации атак на основе поведенческого анализа позволил обеспечить гибкость настройки параметров и возможность удовлетворять требуемые показатели точности идентификации при снижении среднего количества идентификационных признаков. Однако, необходимо исследовать эффективность усовершенствованного метода в условиях постепенного изменения значений признакового пространства при переходе от нормального поведения сети к поведению под атакой.

Представленный метод идентификации атак с использованием параметра степени уверенности показал возможность снизить нагрузку на сети маломощных сенсорных устройств, часто являющихся основой кибер-физических систем. Также, описанный метод может послужить базой для дальнейших разработок и исследований, направленных на адаптацию известных алгоритмов и методов обнаружения вторжений и идентификации атак в беспроводных сенсорных сетях.

ЛИТЕРАТУРА

1. Шелухин О.И., Филинова А. С., Васина А. В. Обнаружение аномальных вторжений в компьютерные сети статистическими методами // T-Comm: Телекоммуникации и транспорт.— 2015.— Том 9.— No10.— С. 42–49.
2. Шилов И. М. Оценка аномального поведения узлов беспроводной сенсорной сети на основе статистических методов // Выпускная квалификационная работа бакалавра.— 2017.— Университет ИТМО
3. Chow T.W.S., Lau B. C.P., Maa E. W.M. Probabilistic fault detector for wireless sensor network. Expert Systems with Applications, (41):3703–3711, 2014.
4. L. Kadam H. Chaudhari. Wireless sensor networks: Security, attacks and challenges. International Journal of Networking, 1(1), 2011.
5. 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
6. 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
7. Shams B., Alrajeh N. A., Khan S. Intrusion detection systems in wireless sensor networks: A review. Inter- national Journal of Distributed Sensor Networks, 9(5), 2013.
8. Shilov I., Korzhuk V., 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
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