ISSN 1992-6502 (Print) Вестник УГАТУ ISSN 2225-2789 (Online)
2015. Т. 19, № 4 (70). С. 95-105. http://journal.ugatu.ac.ru

УДК 004.732:004.056

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОБНАРУЖЕНИЯ АТАК В ЛОКАЛЬНЫХ БЕСПРОВОДНЫХ СЕТЯХ

В. И. ВАСИЛЬЕВ [1], И.В. ШАРАБЫРОВ [2]

[1] vasilyev@ugatu.ac.ru, [2] ilyashar@mail.ru

ФГБОУ ВПО «Уфимский государственный авиационный технический университет» (УГАТУ)

Поступило в редакцию 26.10.2015

Аннотация. На сегодняшний день беспроводные сети передачи данных, в том числе и локального типа, продолжают стремительно развиваться. Однако зачастую безопасность в данных сетях не соответствует необходимому уровню, так как объекты беспроводной инфраструктуры имеют уязвимости и подвержены различным сетевым атакам. Одним из наиболее актуальных средств защиты от беспроводных атак являются системы обнаружения атак. В то же время в связи с широкими возможностями методов интеллектуального анализа данных задачу анализа параметров сетевого трафика на предмет наличия признаков атаки можно решать путем применения данных методов. В статье приведен обзор атак, актуальных для локальных беспроводных сетей, предлагается архитектура системы обнаружения атак на базе методов интеллектуального анализа данных, а также приведено сравнение данных методов при обнаружении рассмотренных типов атак. Результаты экспериментов позволяют сделать вывод о практической значимости предложенного подхода к обнаружению атак в локальных беспроводных сетях.

Ключевые слова. Локальная беспроводная сеть; сетевая атака; модель обнаружения; интеллектуальный анализ данных; сигнатура; Wi-Fi.

ВВЕДЕНИЕ

В настоящее время беспроводные сети завоевали огромную популярность. Согласно прогнозу компании Cisco Systems, к 2017 г. половина всего генерируемого трафика в корпоративных информационных сетях будет приходиться на беспроводные устройства. Это обусловлено, в том числе, и ростом пропускной способности беспроводных сетей. В июне 2013 г. был ратифицирован стандарт 802.11ac, обеспечивающий в перспективе скорость передачи до 6,9 Гбит/с благодаря использованию новых принципов модуляции, поддержке до восьми пространственных потоков передачи данных и полосы шириной до 160 МГц.

Беспроводные сети имеют неоспоримые преимущества перед традиционными кабельными сетями: простота развертывания, мобильность пользователей в зоне действия сети, простое подключение новых пользователей.

С другой стороны, невысокий уровень безопасности таких сетей зачастую ограничивает их применение. В последние годы атаки на локальные беспроводные сети стали обычным явлением. Если при атаке на проводную сеть злоумышленник должен иметь физическое подключение к сети, то в случае беспроводных сетей он может находиться в любой точке зоны действия сети. Кроме того, данные сети подвержены, в том числе по причине несовершенства протоколов, специфическим атакам, которые будут рассмотрены ниже.

Таким образом, можно сформулировать основные проблемы защиты информации в беспроводных сетях:

- распространение сигнала за пределы контролируемой зоны;
- легкий доступ злоумышленника к беспроводному каналу передачи по сравнению с кабельными сетями;
- использование уязвимых протоколов и методов аутентификации;
- отсутствие полноценной защиты от атак со стороны выпускаемых к стандартам дополнений (так протокол 802.11w, утвержденный в 2009 г. и призванный защитить управляющие кадры, обеспечивает их целостность только по-

сле обмена ключами и не распространяется на контрольную информацию);

- ошибки в настройке различных компонентов беспроводной сети.

В связи с вышесказанным исследователи ведут поиск возможных усовершенствований текущих протоколов. Например, в работе [1] автор предлагает шифровать весь блок данных протокола MAC (MPDU), включая MAC-заголовки, кроме последовательности проверки кадра FCS, что, очевидно, приведет к заметным задержкам в передаче данных и низкой пропускной способности канала. Другой подход заключается в помещении в управляющий кадр хэша некой строки, известной только конкретному отправителю, путем передачи которой в дальнейшем его можно однозначно идентифицировать и обработать запрос [2]. Однако данный метод позволяет предотвратить только один вид атаки.

На практике для защиты от сетевых атак рядовые пользователи и небольшие организации, как правило, ограничиваются использованием антивирусного программного обеспечения, которое на современном этапе развития имеет ряд дополнительных модулей защиты (встроенные межсетевые экраны, проверка электронной почты и т. д.). Крупные предприятия вынуждены приобретать дорогостоящие беспроводные системы обнаружения вторжений (Wireless Intrusion Detection Systems, WIDS). Однако в данной области на настоящий момент отсутствуют общепринятые стандарты, производители используют закрытые алгоритмы выявления и классификации атак. При этом задачу отнесения фрагмента сетевого трафика к какому-либо типу атаки или к нормальной сетевой активности можно решать путем применения методов интеллектуального анализа данных (ИАД) [3].

В работах [4, 5] для решения этой задачи предлагается применение нейронных сетей и метода опорных векторов (Support Vector Machine, SVM). В статье [6] приведен вариант комбинации SVM и деревьев принятия решений для обеспечения мультиклассового распознавания атак. В работе [7] рассмотрен подход к организации нейросетевой системы обнаружения атак на базе двухслойного персептрона и сети Кохонена.

Стоит отметить, что приведенные выше исследования российских и зарубежных авторов относятся к обнаружению вторжений в традиционные проводные сети. Однако работы, посвященные целенаправленному применению методов ИАД для обнаружения атак, характерных для локальных беспроводных сетей, в доступной литературе отсутствуют. По этой причине в данной статье рассматриваются основные типы атак, присущие беспроводным сетям, некоторые рекомендуемые способы защиты от них, а также предлагается архитектура системы обнаружения атак на базе методов ИАД и производится оценка эффективности используемых в ней алгоритмов обнаружения атак.

1. АТАКИ НА БЕСПРОВОДНЫЕ СЕТИ

В основе атак на беспроводные сети лежит перехват сетевого трафика от/к точке доступа или трафика между двумя подключенными станциями, а также внедрение дополнительных (поддельных) данных в сеанс беспроводной связи. Для формирования лучшего представления о типах беспроводных атак, которые злоумышленник может осуществить против беспроводной сети, важно их классифицировать. Так, атаки могут быть направлены на разные слои модели OSI: прикладной, транспортный, сетевой, канальный и физический. Специфичными для беспроводных сетей являются физический и канальный уровни, на использовании которых основана группа стандартов IEEE 802.11. Именно использование уязвимостей протоколов и технологий этих уровней является основой проведения атак на локальную беспроводную сеть и первоначальной стадией атак на информационную систему через несанкционированное получение доступа в беспроводную сеть.

В зависимости от цели атаки на локальные беспроводные сети, реализуемые на физическом и канальном уровнях модели OSI, можно поделить на несколько категорий [8]:

- получение несанкционированного доступа к сети:
  - а) ложные точка доступа или клиент (Rogue Access Point);
  - б) подделка MAC-адреса (MAC Spoofing);
  - в) взлом клиента сети;
  - г) взлом точки доступа.
- нарушение целостности:
  - а) инъекция поддельных кадров (802.11 Frame Injection);
  - б) повтор, удаление пакетов с данными (802.11 Data Replay, 802.11 Data Deletion);
  - в) перехват и воспроизведение пакетов EAP, RADIUS (802.1X EAP Replay, 802.1X RADIUS Replay);
- нарушение конфиденциальности:
  - а) подслушивание (Eavesdropping);
  - б) атака «злой двойник» (Evil Twin);

в) фишинг с помощью ложной точки доступа (AP Phishing);

г) атака «человек посередине» (Man in the Middle);

- нарушение доступности:

а) радиочастотное зашумление;

б) захват среды передачи ложными RTS/CTS-кадрами (Queensland DoS);

в) наводнение запросами на подключение (Probe Request Flood);

г) наводнение кадрами ассоциации, аутентификации, диссоциации, деаутентификации (Associate / Authenticate / Disassociate / Deauthenticate Flood);

д) наводнение кадрами EAP (802.1X EAPStart, EAPFailure Flood);

- обход процедуры аутентификации:

а) взлом Pre-Shared Key;

б) кража личности 802.1X (802.1X Identity Theft);

в) понижение уровня безопасности EAP (802.1X EAP Downgrade);

г) взлом пароля 802.1X;

д) взлом доменных учетных записей;

е) взлом WPS PIN.

Данные атаки основаны на эксплуатации уязвимостей беспроводных сетей, представленных в базе WVE [9]:

- отправка Probe-запросов с полем тега SSID нулевой длины (WVE-2006-0064);
- EAP Logoff атака (WVE-2005-0050);
- RTS/CTS наводнение (WVE-2005-0051);
- наводнение WLAN пакетами диссоциации (WVE-2005-0046);
- наводнение WLAN пакетами деаутентификации (WVE-2005-0045);
- KARMA framework (WVE-2006-0032);
- отправка неверного кода причины деаутентификации;
- отправка слишком длинного SSID (WVE-2006-0071, WVE-2007-0001);
- отправка Airjack beacon кадра (WVE-2005-0018);
- отправка неверного номера канала в beacon кадрах (WVE-2006-0050).

Для формирования базы беспроводных атак канального уровня была организована тестовая локальная беспроводная сеть с технологией защиты доступа WPA2-Enterprise. Средой генерации атак служил ноутбук с установленным дистрибутивом Kali Linux версии 1.1.0 с набором специальных утилит для тестирования на проникновение в сеть и беспроводным адаптером Atheros AR9485 в режиме мониторинга. Для перехвата и анализа кадров с целью формирования базы сигнатур атак использовался второй ноутбук с Windows 7 и беспроводным адаптером Atheros AR9285 в пассивном режиме мониторинга. Собранные кадры были проанализированы: каждый кадр был обозначен как нормальный либо как являющийся частью какого-либо типа атаки, относящегося к одной из следующих четырех категорий атак:

- получение несанкционированного доступа к сети (Access Control);
- нарушение целостности (Integrity Violation);
- нарушение конфиденциальности (Confidentiality Violation);
- нарушение доступности (Denial of Service).

Соотношение числа атак данных категорий в базе сигнатур показано в табл. 1.

Таблица 1

Соотношение количества сигнатур атак канального уровня в обучающей и тестовой базе

Обучающая база		*Тестовая база*	
Класс	Кол-во	Класс	Кол-во
Normal	114959	Normal	27886
Access Control		Access Control	
rogue_client	131	rogue_client	225
MAC_spoofing	106	MAC_spoofing	12
fake_auth	81	fake_auth	9
caffe_latte	198	caffe_latte	22
chopchop	202	chopchop	22
client_fragment	56862	client_fragment	6318
AP_fragment	3552	AP_fragment	395
Confidentiality Violation		Confidentiality Violation	
evil_twin_AP	1344	evil_twin_AP	149
Integrity Violation		Integrity Violation	
data_replay	75785	data_replay	8420
EAP_replay	75	EAP_replay	8
Denial of Service		Denial of Service	
beacon_flood	1761	beacon_flood	2153
auth_flood	569	auth_flood	74
deauth_flood	10940	deauth_flood	1216
EAPOL_start_ flood	16345	EAPOL_start_ flood	1816
EAPOL_logoff_ flood	1923	EAPOL_logoff_ flood	2425
RTS/CTS_flood	1831	RTS/CTS_flood	2237

В качестве образцов атак с сетевого по прикладной уровни использована усовершенствованная база сигнатур NSL KDD-2009 [10], построенная на основе базы KDD-99 по инициативе американской Ассоциации оборонных научных исследований (DARPA) [11]. Целью проекта являлась оценка эффективности исследований в области обнаружения вторжений. В результате был собран набор данных о соедине-

ниях, который охватывает широкий спектр различных вторжений, смоделированных в среде локальной сети, имитирующей сеть Военно-воздушных сил США. Данные собирались в течение девяти недель.

Соединение представляет собой последовательность пакетов, начинающуюся и заканчивающуюся в определенные моменты времени, между которыми потоки данных передаются от IP-адреса источника к IP-адресу получателя по определенному протоколу. Каждое соединение обозначено как нормальное либо как какой-то тип атаки одной из четырех категорий: отказ в обслуживании (Denial of Service, DoS), несанкционированное получение прав пользователя (Remote to Local, R2L), несанкционированное повышение прав пользователя до суперпользователя (User to Root, U2R) и зондирование (Probe). Подробное описание этих типов атак приведено в [12]. Соотношение числа атак разных типов показано в табл. 2 и 3.

Таблица 2

Соотношение количества сигнатур атак в обучающей базе NSL KDD-2009

Normal	67343		
DoS		R2L	
Класс	Кол-во	Класс	Кол-во
neptune	41214	guess_passwd	162
smurf	2646	ftp_write	8
Pod	201	imap	11
teardrop	892	phf	4
land	18	multihop	7
back	956	warezmaster	20
U2R		Probe	
Класс	Кол-во	Класс	Кол-во
buffer_overflow	30	portsweep	2931
loadmodule	9	ipsweep	3599
perl	3	satan	3633
rootkit	10	nmap	1493

Некоторые из указанных выше типов атак являются издержками самой технологии радиочастотной передачи данных (глушение), а также зависят от человеческого фактора и должны решаться с помощью организационных мер. Среди технических средств защиты сети, помимо межсетевых экранов, списков контроля доступа и других традиционных средств, следует выделить беспроводные системы обнаружения атак.

Таблица 3

Соотношение количества сигнатур атак в тестовой базе NSL KDD-2009

Normal	9711		
DoS		R2L	
Класс	Кол-во	Класс	Кол-во
neptune	4657	guess_passwd	1231
smurf	665	ftp_write	3
Pod	41	imap	1
teardrop	12	phf	2
land	7	multihop	18
back	359	warezmaster	944
U2R		Probe	
Класс	Кол-во	Класс	Кол-во
buffer_overflow	20	portsweep	157
loadmodule	2	ipsweep	141
perl	2	satan	735
rootkit	13	nmap	73

2. СИСТЕМА ОБНАРУЖЕНИЯ АТАК

В отличие от традиционных систем обнаружения атак, получающих все пакеты сети, беспроводные системы производят выборку сетевого трафика. Стандарты семейства 802.11 используют два основных диапазона частот: 2,4 ГГц и 5 ГГц, которые, в свою очередь, делятся на каналы. WIDS обеспечивают поочередное сканирование каналов на предмет наличия активных атак. Принятие решения о безопасности какой-либо сетевой активности в коммерческих продуктах реализуется с помощью закрытых алгоритмов, принцип работы которых составляет коммерческую тайну. При этом заявленное количество и виды обнаруживаемых атак у разных продуктов отличаются, хотя в действительности они принадлежат одному типу атак, что объясняется отсутствием стандартов в области беспроводных атак.

Как было показано в упомянутых выше работах, задачи обнаружения и классификации атак можно решать с помощью применения методов ИАД, позволяющих выявить значимые корреляции, образцы и тенденции в больших объемах данных. В предлагаемой системе используются алгоритмы построения классифицирующей модели на базе метода опорных векторов, метода k-ближайших соседей, нейронных сетей и деревьев принятия решений.

Предлагаемая архитектура интеллектуальной системы обнаружения атак обладает модульной схемой организации взаимодействия между компонентами с выделенной подсистемой сенсоров и централизованным управлением через консоль администратора. Архитектура системы представлена на рис. 1.

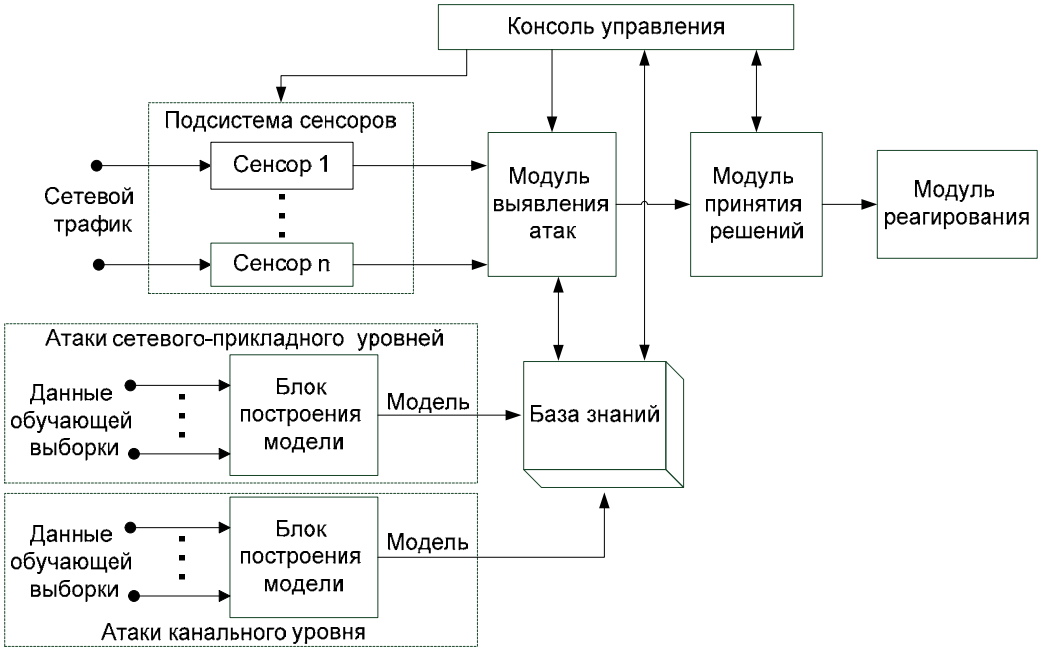

Рис. 1. Архитектура системы обнаружения атак

Основой для выявления атак является база знаний, построение которой на этапе начального конфигурирования системы обеспечивает блок построения классифицирующей модели. Классифицирующая модель строится на основе сигнатур обучающей выборки и затем используется для классификации реальной сетевой активности.

Модуль выявления атак проектируемой системы обнаружения атак функционально можно разделить на:

- подмодуль обнаружения атак сетевого, транспортного и прикладного уровней;
- подмодуль обнаружения атак канального уровня.

Система работает в двух режимах:

- режим конфигурирования (обучения), когда в качестве входных данных в блок построения классифицирующей модели загружается набор сигнатур, каждая из которых представляет собой пару {вектор параметров трафика | тип атаки}. На основе этого набора строится классифицирующая модель;
- режим нормальной работы, когда значения параметров трафика подаются в качестве входных данных на подсистему сенсоров. Далее модуль выявления атак с помощью построенной на предыдущем этапе классифицирующей модели определяет, соответствуют ли показания сенсоров нормальному состоянию либо той или иной атаке, и передает результат подсистеме принятия решений, которая, в случае атаки или подозрительной активности, выдает оповещение на консоль управления (пассивное обнаружение) и формирует команду модулю реагирования (активное блокирование).

Далее мы подробнее рассмотрим методы ИАД, составляющие основу алгоритма построения классифицирующей модели предлагаемой системы.

3. МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Метод опорных векторов (SVM) относится к методам линейной классификации. Каждое состояние системы представляется в виде точки в многомерном пространстве, координатами которого являются характеристики системы. Два множества точек, принадлежащих к двум разным классам, разделяются гиперплоскостью в этом пространстве. При этом гиперплоскость строится так, чтобы расстояния от нее до ближайших экземпляров обоих классов (опорных векторов) были максимальны, что обеспечивает наибольшую точность классификации.

На рис. 2 приведен пример классификации объектов в двумерном пространстве с помощью SVM.

На рисунке представлен обучающий набор данных, представляющий собой множество точек вида $\{x_i, y_i\}$, $i = 1,..., l$, где $x_i \in \mathcal{R}^n$, а $y_i \in \{1, -1\}$ – индикатор класса, к которому относится точка x_i. Классы точек линейно разделимы, т. е. существует такая гиперплоскость, по одну сторону которой лежат точки класса $y_i = 1$, а по другую –

класса $y_i = -1$. Точки, расположенные непосредственно на гиперплоскости, удовлетворяют уравнению:

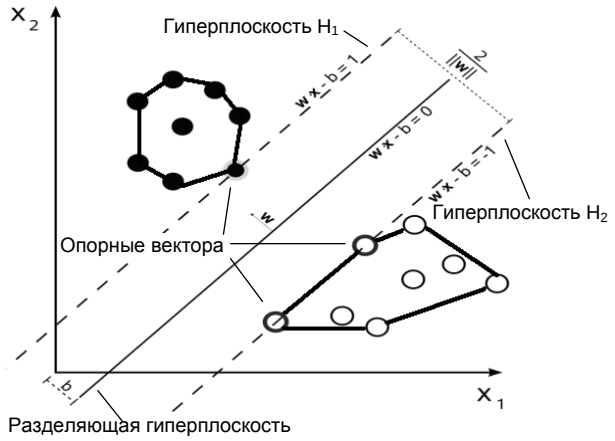

Рис. 2. Пример классификации SVM

$$w \bullet x - b = 0 \ , \qquad (1)$$

где вектор w – перпендикуляр к разделяющей гиперплоскости, величина $|b|/||w||$ (абсолютная величина b, деленная на модуль вектора w) определяет расстояние от начала координат до гиперплоскости, оператор • обозначает скалярное произведение в евклидовом пространстве, в котором лежат данные.

Все точки, для которых выполняется условие $w \bullet x_i - b = 1$, лежат в гиперплоскости H_1, параллельной к разделяющей гиперплоскости, и на расстоянии $|1 - b|/||w||$ от начала координат. Похожим образом те точки, для которых выполняется условие $w \bullet x_i - b = -1$, лежат в гиперплоскости H_2, параллельной к плоскости H_1 и к разделяющей гиперплоскости, на расстоянии $|-1 - b|/||w||$ от начала координат. Таким образом, расстояние между плоскостью и положительным (отрицательным) опорным вектором равно $1/||w||$, и, следовательно, ширина полосы (зазора) равна $2/||w||$.

В качестве достоинств данного метода можно выделить высокую точность, способность к обобщению и низкую вычислительную сложность принятия решения. Недостатком является относительно большая вычислительная сложность построения классифицирующей модели.

В работах [13, 14] исследуется способ обнаружения атак на основе метода опорных векторов. Метод использовался для построения классифицирующей модели из данных обучающей выборки. Модель опробована на атаках типа переполнение буфера, руткит и SYN-наводнение и показала актуальность применения метода опорных векторов в качестве основы системы обнаружения атак.

Метод k-ближайших соседей (k-nearest neighbor, k-NN) - метод классификации, основным принципом которого является присваивание объекту того класса, который является наиболее распространенным среди соседей данного объекта. Соседи образуются из множества объектов, классы которых уже известны, и, исходя из заданного значения k ($k \geq 1$), определяется, какой из классов наиболее многочислен среди них. Если $k = 1$, то объект просто относится к классу единственного ближайшего соседа.

Метод k-NN является одним из самых простых методов ИАД. Недостатком метода k-NN является то, что он чувствителен к локальной структуре данных.

Нейронные сети позволяют решать практические задачи, связанные с распознаванием и классификацией образов. Нейронная сеть состоит из взаимосвязанных нейронов, образующих входной, промежуточные (скрытые) и выходной слои. Обучение происходит путем корректировки значений весов нейронов для минимизации ошибки классификации.

Преимуществами нейронных сетей являются их способность автоматически приобретать знания в процессе обучения, а также способность к обобщению, основной недостаток – чувствительность к шуму во входных данных.

Деревья принятия решений представляют собой древовидную структуру из «листьев» и «ветвей». На ребрах («ветвях») дерева принятия решений записаны атрибуты, от которых зависит целевая функция, в «листьях» записаны значения целевой функции, а в остальных узлах – атрибуты, по которым различаются объекты. Чтобы классифицировать новый объект, надо спуститься по дереву от корня до листа и получить соответствующий класс, т. е. путь от корня до листа выступает правилами классификации на основе значений атрибутов объекта.

Достоинства деревьев принятия решений – простой принцип их построения, хорошая интерпретируемость результатов, недостаток – невысокая точность классификации.

Далее для выявления наиболее эффективного метода построения классифицирующей модели применительно к беспроводной системе обнаружения атак будет приведено сравнение рассмотренных методов ИАД.

4. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Оценка корректности распознавания рассмотренных типов атак с помощью СОА произведена путем сравнения результатов классификации с помощью различных методов ИАД.

На основании построенной выше классификации атак по уровням модели OSI атаки на локальные беспроводные сети разделены на две группы:

- атаки физического и канального уровня, являющиеся специфичными для беспроводных сетей;
- атаки с сетевого по прикладной уровни, присущие любой технологии организации локальных сетей, в том числе Ethernet.

В качестве образцов атак сетевого и прикладного уровней соответствующий подмодуль обнаружения атак предлагаемой системы в ходе экспериментов использует сигнатуры базы NSL KDD-2009.

Изначально для описания атак в базе NSL-KDD-2009 использован 41 признак, отражающий прикладной, транспортный и сетевой уровни модели OSI. Однако часть предлагаемых признаков не применима для современных сетевых атак по причине своей неактуальности [15], в связи с чем количество признаков было сокращено. Выбранные признаки представлены в табл. 4. Параметры трафика, использованные для описания сигнатур атак канального уровня, указаны в табл. 5. Для описания атак, характеризующихся большим количеством соединений к узлу назначения, было выбрано окно длительностью две секунды (атаки DoS), а также окно в 100 соединений с одним и тем же узлом (Probe).

Таблица 4

Значимые параметры трафика для сетевого-прикладного уровней

Характеристика	*Описание*	*Тип*
Характеристики TCP-соединения		
duration	Продолжительность соединения (с)	численный
protocol_type	Протокол транспортного уровня	текстовый
service	Сервис прикладного уровня	текстовый
flag	Статус соединения	бинарный
src_bytes	Входящий поток, байт	численный
dst_bytes	Исходящий поток, байт	численный
land	Адреса совпадают, 0 иначе	бинарный
wrong_fragment	Число неправильных фрагментов	численный
urgent	Число срочных пакетов	численный
Характеристики сеанса		
hot	Число «горячих» индикаторов	численный
num_failed_logins	Число неудачных попыток входа	численный
logged_in	Успешный вход	бинарный
root_shell	Доступ с административными полномочиями	бинарный
num_root	Число попыток доступа с административными полномочиями	численный
num_shells	Число попыток использования командной строки	численный
num_access_files	Число операций с файлами контроля доступа	численный
Статистика за 2 секунды / за 100 соединений		
count / dst_host_count	Число соединений с совпадающим хостом	численный
serror_rate/ dst_host_serror_rate	% соединения с ошибкой ``SYN''	численный
rerror_rate / dst_host_same_src_port_rate	% соединений с ошибкой ``REJ'' / % соединений с одинаковым исходным портом	численный
same_srv_rate / dst_host_same_srv_rate	% соединений с одинаковым сервисом	численный
diff_srv_rate / dst_host_diff_srv_rate	% соединений с различным сервисом	численный
srv_count / dst_host_srv_count	Число соединений с совпадающим сервисом	численный
srv_serror_rate / dst_host_srv_serror_ rate	% соединений с ошибкой ``SYN''	численный
srv_rerror_rate / dst_host_srv_rerror_ rate	% соединений с ошибкой ``REJ''	численный
srv_diff_host_rate / dst_host_srv_diff_host_ rate	% соединений с различающимися хостами	численный

Таблица 5

Значимые параметры трафика для канального уровня

Характеристика	*Описание*	*Тип*
Характеристики протоколов 802.11		
frame_ type/subtype	Тип/подтип кадра	текстовый
protocol_type	Тип протокола канального уровня	текстовый
source_address	MAC-адрес источника	текстовый
destination_address	MAC-адрес назначения	текстовый

Окончание табл. 5

length	Размер кадра, байт	численный
SSID	Значение тега SSID	текстовый
sequence_number	Номер кадра	численный
fragment_number	Номер фрагмента	численный
DS_status	Участие распределенной системы в обмене	численный
more_fragments	Еще фрагменты для передачи, 0 иначе	бинарный
retry	Повторная передача предыдущего кадра, 0 иначе	бинарный
pwr_mgt	Клиент в режиме энергосбережения, 0 иначе	бинарный
more_data	Буферизованные кадры для передачи, 0 иначе	бинарный
protected_flag	Данные кадра зашифрованы, 0 иначе	бинарный
order_flag	Обработка кадров строго по порядку, 0 иначе	бинарный
duration	Продолжительность передачи ACK+SIFS, мкс	численный
chan_number	Номер канала	численный
signal	Уровень сигнала передатчика, %	численный
TX_rate	Скорость передачи, Мбит/с	численный
cipher	Используемый алгоритм шифрования	текстовый
reason_code	Код причины деаутентификации	численный
Статистика за 2 секунды		
mng_frm_count	Число управляющих кадров	численный
ctrl_frm_count	Число контрольных кадров	численный
probe_count	Число запросов на подключение	численный
frag_count	Среднее число фрагментированных пакетов	численный

Эксперименты проводились в среде RapidMiner версии 5.3.015 [16] по схеме, приведенной на рис. 3.

На первом шаге происходила обработка данных из базы, так как для безошибочного функционирования алгоритмов все атрибуты должны иметь численные значения, распределенные между нулем и единицей. Для этого текстовые атрибуты преобразовывались в бинарные, численные нормализовывались относительно минимального и максимального значений.

После этого данные обучающей выборки поступали на вход блока построения классифицирующей модели, составляющей основу базы знаний, различными методами ИАД. Затем модуль выявления атак проводил классификацию

Рис. 3. Схема проведения эксперимента

записей тестовой базы на основании соответствующей модели по критериям, содержащимся в базе знаний, и присваивал метку класса сетевой активности. На основании совпадения предполагаемых и действительных меток классов оценивалась эффективность обнаружения атак по следующим критериям:

1. Общий процент корректно классифицированных атак *A (accuracy)*:

$$A = \frac{TP + TN}{N}, \qquad (2)$$

где *TP* и *TN* – общее количество истинных записей, *N* – общее количество классифицированных записей.

2. Точность классификации *P (precision)*:

$$P = \frac{TP}{TP + FP}, \qquad (3)$$

где *TP* – количество истинно-положительных записей, *FP* – количество ложно-положительных записей.

3. Полнота классификации *R (recall)*:

$$R = \frac{TP}{TP + FN}, \qquad (4)$$

где *FN* – количество ложно-отрицательных записей.

Результаты классификации с помощью различных методов ИАД указаны в табл. 6 и 7.

Метод опорных векторов был реализован с помощью SVM C-SVC библиотеки LibSVM, в качестве функции ядра использовалась радиальная базисная функция (RBF). Величина максимальной ошибки обучения была ограничена значением 10^{-5}.

При классификации методом k-ближайших соседей экспериментальным путем в качестве оптимальных параметров работы алгоритма были выбраны значение *k*, равное пяти, и метрика – Манхэттенское расстояние.

Нейронная сеть была реализована в виде многослойного персептрона с двумя скрытыми слоями. Обучение продолжительностью 1500 циклов производилось с помощью алгоритма обратного распространения ошибки. Величина максимальной ошибки обучения была равна 10^{-7}.

Построение деревьев принятия решений производилось с помощью стандартного оператора среды RapidMiner, минимальный порог для образования нового узла выбирался равным четырем, минимальное количество листьев узла – один, максимальное количество уровней – 10.

Таблица 6

Показатели эффективности определения атак сетевого-прикладного уровней

Группа	Класс сетевой активности	Метод опорных векторов		k-ближайших соседей		Нейронная сеть		Деревья принятия решений	
		Полнота	Точность	Полнота	Точность	Полнота	Точность	Полнота	Точность
DoS	neptune	98,97%	99,98%	97,25%	97,50%	99,36%	99,98%	97,32%	99,93%
normal	normal	96,56%	92,28%	96,55%	93,63%	97,07%	87,25%	97,10%	90,98%
R2L	guess_passwd	76,69%	100,00%	66,86%	95,48%	66,37%	97,03%	65,72%	99,88%
DoS	smurf	100,00%	99,70%	97,59%	100,00%	95,19%	99,53%	100,00%	100,00%
Probe	satan	93,74%	76,47%	94,83%	76,76%	90,75%	81,84%	96,19%	80,62%
U2R	buffer_overflow	25,00%	62,50%	35,00%	100,00%	0,00%	0,00%	25,00%	62,50%
DoS	back	98,05%	98,60%	99,44%	100,00%	96,10%	97,73%	77,16%	92,33%
R2L	warezmaster	59,11%	99,11%	82,20%	99,74%	16,10%	98,06%	63,56%	100,00%
DoS	pod	95,12%	72,22%	95,12%	72,22%	82,93%	70,83%	95,12%	46,99%
Probe	nmap	98,63%	93,51%	97,26%	91,03%	79,45%	90,62%	98,63%	74,23%
Probe	ipsweep	97,16%	93,84%	97,16%	74,86%	97,87%	79,31%	99,29%	88,05%
probe	portsweep	91,08%	56,30%	85,35%	73,22%	89,17%	61,67%	84,71%	54,07%
DoS	teardrop	83,33%	21,28%	83,33%	14,08%	75,00%	18,75%	100,00%	24,49%
DoS	land	57,14%	100,00%	57,14%	100,00%	0,00%	0,00%	14,29%	100,00%
Средняя		**83,61%**	**83,27%**	**84,65%**	**84,89%**	**70,38%**	**70,19%**	**79,58%**	**79,58%**

Таблица 7

Показатели эффективности определения атак канального уровня

Класс	Метод опорных векторов		k-ближайших соседей		Нейронная сеть		Дерево принятия решений	
	Полнота	Точность	Полнота	Точность	Полнота	Точность	Полнота	Точность
normal	98,03%	92,49%	97,65%	99,26%	94,37%	99,38%	95,48%	95,11%
rogue_client	100,00%	37,56%	6,22%	20,00%	32,44%	20,00%	100,00%	69,02%
EAPOL_logoff_flood	8,82%	100,00%	26,85%	100,00%	0,12%	100,00%	44,08%	100,00%
auth_flood	85,14%	94,03%	100,00%	93,67%	100,00%	92,50%	97,30%	100,00%
EAPOL_start_flood	100,00%	100,00%	100,00%	50,58%	100,00%	44,14%	100,00%	100,00%
deauth_flood	100,00%	99,10%	100,00%	99,75%	100,00%	84,39%	100,00%	100,00%
caffe_latte	0,00%	0,00%	100,00%	100,00%	100,00%	70,97%	100,00%	100,00%
chopchop	100,00%	62,86%	100,00%	100,00%	100,00%	3,28%	100,00%	2,27%
client_fragment	97,44%	99,77%	100,00%	99,89%	100,00%	96,98%	100,00%	100,00%
AP_fragment	98,73%	97,01%	99,75%	98,25%	100,00%	98,26%	100,00%	100,00%
data_replay	99,82%	98,13%	100,00%	99,98%	99,96%	99,53%	100,00%	100,00%
MAC_spoofing	100,00%	6,63%	100,00%	10,91%	0,00%	0,00%	0,00%	0,00%
evil_twin_AP	100,00%	100,00%	100,00%	64,78%	100,00%	94,30%	100,00%	94,90%
EAP_replay	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
beacon_flood	100,00%	100,00%	100,00%	99,95%	99,91%	100,00%	100,00%	99,86%
RTS/CTS_flood	99,82%	99,82%	100,00%	84,64%	100,00%	91,49%	100,00%	91,68%
fake_auth	55,56%	100,00%	66,67%	85,71%	77,78%	10,45%	100,00%	100,00%
Средняя	**84,90%**	**81,61%**	**88,07%**	**82,79%**	**82,62%**	**70,92%**	**90,40%**	**85,46%**

Как видно из таблицы 6, методы опорных векторов и k-ближайших соседей показали близкие результаты в ходе обнаружения атак, несколько хуже проявили себя дерево принятия решений и нейронная сеть. Низкий процент обнаружения некоторых типов атак, таких как warezmaster, guess_passwd, buffer_overflow и land, вызван неравномерным количественным распределением образцов обучающей выборки для разных классов – преобладанием нормальных сигнатур и атак категорий DoS и Probe. По этой же причине часть атак была классифицированы неверно, поэтому результаты по ним не представлены в табл. 6. Однако, согласно показателям в табл. 7, метод k-ближайших соседей и дерево принятия решений превосходят SVM и нейронные сети в решении задачи по обнаружению атак канального уровня.

Таким образом, анализ экспериментальных данных показывает, что использованные алгоритмы демонстрируют различные значения показателей эффективности обнаружения атак в зависимости от типа сетевой активности и уровня модели OSI, на котором реализуется атака. В связи с этим предлагается использовать ансамбль из четырех разработанных алгоритмов и одного арбитра, определяющего итоговый класс сетевой активности методом взвешенного голосования. Архитектура и принципы функционирования предлагаемого ансамбля составит суть дальнейших исследований.

ЗАКЛЮЧЕНИЕ

В данной статье представлен обзор сетевых атак, актуальных для локальных беспроводных сетей, представлена архитектура предлагаемой системы обнаружения атак, базирующейся на применении методов ИАД для распознавания данных атак, приведено сравнение данных методов в ходе экспериментов по обнаружению рассмотренных типов атак.

В целом методы показали высокие точность и полноту обнаружения в ходе проведения экспериментов, из чего можно сделать вывод о практической значимости предложенного подхода к обнаружению атак в локальных беспроводных сетях.

Дальнейшие исследования предполагается продолжить в направлении исследования новых типов атак в локальных беспроводных сетях, а также организации модульной структуры системы обнаружения вторжений, с использованием ансамбля рассмотренных в данной статье методов ИАД.

СПИСОК ЛИТЕРАТУРЫ

1. **Ross D.** Securing IEEE802.11 Wireless LANs. PhD thesis, Queensland University of Technology, 2010 [Электронный ресурс]. URL: http://eprints.qut.edu.au/37638/1/David_Ross_Thesis.pdf (дата обращения 28.01.2013). [D. Ross (2013, Jan. 28). *Securing IEEE802.11 Wireless LANs.* PhD thesis, Queensland University of Technology, 2010 [Online]. Available: http://eprints.qut.edu.au/37638/1/David_Ross_Thesis.pdf]

2. **Nguyen T., Nguyen B., Pham H.** An efficient solution for preventing Dis'ing attack on 802.11 networks // The 2012 International Conference on Green Technology and Sustainable Development (GTSD2012): Journal of Engineering Technology and Education, Hochiminh City, 2012. P. 395–403. [T. Nguyen, B. Nguyen and H. Pham, "An efficient solution for preventing Dis'ing attack on 802.11 networks", The 2012 International Conference on Green Technology and Sustainable Development (GTSD2012), in *Journal of Engineering Technology and Education*, Hochiminh, 2012, pp. 395-403.]

3. **Sinclair C., Pierce L., Matzner S.** An Application of Machine Learning to Network Intrusion Detection // Proceedings of Computer Security Applications Conference, Dec. 6–10 1999. (ACSAC '99). P. 371–377. [C. Sinclair, L. Pierce and S. Matzner, "An Application of Machine Learning to Network Intrusion Detection", in *Proc. of Computer Security Applications Conference*, Dec. 6-10 1999 (ACSAC '99), pp. 371-377.]

4. **Tang H., Cao Z.** Machine Learning-based Intrusion Detection Algorithms // Journal of Computational Information Systems, 2009. P. 1825–1831. [H. Tang and Z. Cao, "Machine Learning-based Intrusion Detection Algorithms", in *Journal of Computational Information Systems*, 2009, pp. 1825-1831.]

5. **Mukkamala S., Janoski G., Sung A.** Intrusion Detection: Support Vector Machines and Neural Networks [Электронный ресурс]. URL: http://www.cs.uiuc.edu/class/fa05/cs591han/papers/mukkCNN02.pdf (дата обращения 09.01.2013). [S. Mukkamala, G. Janoski and A. Sung (2013, Jan. 09), *Intrusion Detection: Support Vector Machines and Neural Networks* [Online]. Available: http://www.cs.uiuc.edu/class/fa05/cs591han/papers/mukkCNN02.pdf.]

6. **Mulay S., Devale P., Garje G.** Intrusion Detection System using Support Vector Machine and Decision Tree // International Journal of Computer Applications, June 2010. V. 3, N.3. P. 40–43. [S. Mulay, P. Devale and G. Garje, "Intrusion Detection System using Support Vector Machine and Decision Tree", in *International Journal of Computer Applications*, 2010, vol. 3, no. 3, pp. 40-43.]

7. **Нейросетевая** технология обнаружения сетевых атак на информационные ресурсы / Ю. Г. Емельянова [и др.] // Программные системы: теория и приложения: электронный научный журнал. 2011. № 3(7). С. 3–15. [J. G. Emelyanova, *et al.*, "Neural network technology of detection network attacks on information resource", (in Russian), in *Programmnyye sistemy: teoriya i prilozheniya*: electronic scientific journal, no. 3(7), pp. 3-15, 2011.]

8. **Arinze N.** Wireless Local Area Network (WLAN): Security Risk Assessment and Countermeasures. Blekinge Institute of Technology, 2008 [Электронный ресурс]. URL: http://www.bth.se/fou/cuppsats.nsf/all/2cf7d7f61e47ae4ec1257514004fce3f/$file/WLAN_Security%20Risk%20Assessment%20and%20Countermeasures.pdf (дата обращения 12.03.2013). [N. Arinze (2013, Mar. 12). *Wireless Local Area Network (WLAN): Security Risk Assessment and Countermeasures.*

Blekinge Institute of Technology, 2008 [Online]. Available: http://
www.bth.se/fou/cuppsats.nsf/all/2cf7d7f61e47ae4ec1257514004fce3f/$file/WLAN_Security%20Risk%20Assessment%20and%20Countermeasures.pdf]

9. **WVE.** Wireless Vulnerabilities and Exploits [Электронный ресурс]. URL: http://www.wve.org (дата обращения 05.10.2013). [WVE (2013, Oct. 05). *Wireless Vulnerabilities and Exploits* [Online]. Available: http://www.wve.org]

10. **The NSL-KDD** Data Set. [Электронный ресурс]. URL: http://nsl.cs.unb.ca/NSL-KDD (дата обращения 22.01.2013). [The NSL-KDD Data Set (2013, Jan. 22) [Online]. Available: http://nsl.cs.unb.ca/NSL-KDD]

11. **KDD cup 99** Intrusion detection data set. [Электронный ресурс]. URL: http://kdd.ics.uci.edu/databases/kddcup99/kddcup.data.gz (дата обращения 19.11.2011). [KDD cup 99 Intrusion detection data set (2011, Nov. 19) [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup.data.gz]

12. **Lincoln Laboratory.** DARPA Intrusion Detection Evaluation. [Электронный ресурс]. URL: http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/docs/attackDB.html (дата обращения 01.04.2012). [Lincoln Laboratory. *DARPA Intrusion Detection Evaluation.* (2012, Apr. 01) [Online]. Available: http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/docs/ attackDB.html]

13. **Миронов К. В., Шарабыров И. В.** О применении метода опорных векторов в системах обнаружения атак // Мавлютовские чтения: Всероссийская молодежная научная конференция: сб. тр. Т. 3. Уфа, УГАТУ, 2012. С. 28–30. [K. V. Mironov and I. V. Sharabyrov, “Application of support vector machine in intrusion detection systems”, (in Russian), in *Proc. of National Youth Scientific Conference “Mavlyutov readings”*, UGATU, vol. 3, pp. 28-30, 2012.]

14. **Разработка** модели обнаружения сигнатур атак на основе метода опорных векторов / В.И. Васильев [и др.] // Материалы XII Международной научно-практической конференции «Информационная безопасность-2012». Ч.1. Таганрог: Изд-во ТТИ ЮФУ, 2012. С. 192–201. [V. I. Vasilyev, *et al.,* “Development of attacks signature detection model on the base of support vector machine”, in *Proc. of XII International scientific and practical conference “Information Security 2012”*, Taganrog, 2012, vol. 1, pp. 192-201.]

15. **Olusola A., Oladele A., Abosede D.** Analysis of KDD ’99 Intrusion Detection Dataset for Selection of Relevance Features // Proceedings of the World Congress on Engineering and Computer Science, (San Francisco, Oct. 20–22 2010). V. 1. P. 162–168. [A. Olusola, A. Oladele and D. Abosede. “Analysis of KDD ’99 Intrusion Detection Dataset for Selection of Relevance Features”, in *Proc. of the World Congress on Engineering and Computer Science*, San Francisco, 2010, vol. 1, pp. 162-168.]

16. **RapidMiner Studio.** [Электронный ресурс]. URL: https://rapidminer.com (дата обращения 01.09.2013). [RapidMiner Studio (2013, Sept. 01). [Online]. Available: https://rapidminer.com]

ОБ АВТОРАХ

ВАСИЛЬЕВ Владимир Иванович, проф., зав. каф. выч. техн. и защиты инф. Дипл. по спец. «Промышленная электроника» (УГАТУ, 1970). Д-р техн. наук (УГАТУ, 1989). Иссл. в обл. искусственного интеллекта, защиты инф., интеллектуального упр. в сложных техн. и организационных системах.

ШАРАБЫРОВ Илья Викторович, асп. каф. выч. техн. и защиты инф. Дипл. спец. по защите инф. (УГАТУ, 2012). Готовит дисс. в обл. обнаружения атак в локальных беспроводных сетях на основе методов интеллектуального анализа данных.

METADATA

Title: Intelligent intrusion detection system in local wireless networks.

Authors: V. I. Vasilyev[1], I. V. Sharabyrov[2]

Affiliation:
Ufa State Aviation Technical University (USATU), Russia.

Email:
[1] vasilyev@ugatu.ac.ru.
[2] ilyashar@mail.ru.

Language: Russian.

Source: Vestnik UGATU (scientific journal of Ufa State Aviation Technical University), V. 19, N. 4 (70), P. 95–105, 2015. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

Abstract: Nowadays wireless networks, including local ones, continue to evolve rapidly. However security in these networks does not often correspond to the required level, as objects of wireless infrastructure have vulnerabilities and are susceptible to various network attacks. One of the most actual protection means from the wireless attacks are the intrusion detection systems. At the same time, due to the wide possibilities of data mining methods the task of network traffic parameters analysis for the signs of attack can be solved by application of these methods. The article provides an overview of network attacks that are relevant to local wireless networks, the proposed architecture of intrusion detection system based on data mining techniques, as well as a comparison of these methods in detecting the types of attacks mentioned above. The experimental results allow making the conclusion about the practical relevance of the proposed approach for intrusion detection in local wireless networks.

Key words: Local wireless network; network attack; detection model; data mining; signature; Wi-Fi.

About authors:

VASILYEV, Vladimir Ivanovich, Prof., Head of the Dept. of Computer Engineering and Information Protection. Dipl. specialist "Industrial Electronics" (UGATU, 1970). Cand. of Tech. Sci. (UGATU, 1975), Dr. of Tech. Sci. (UGATU, 1989). Researches in the field of artificial intelligence, information protection, intelligent control in complex technical and organizational systems.

SHARABYROV, Ilya Viktorovich, Postgrad. (PhD) Student, Dept. of Computer Engineering and Information Protection. Dipl. specialist of Information Protection (UGATU, 2012). Prepares a thesis “Intelligent intrusion detection system in local wireless networks”.