БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Цыбулин А.М., Свищева М.Н.* Системный подход к повышению эффективности борьбы с инсайдерской деятельностью пользователей информационной системы организации // Известия ЮФУ. Технические науки. – 2013. – № 12 (149). – С. 25-33.
2. *Бешта А.А.* Архитектура агента контроля над внутренним злоумышленником на основе механизма оценки доверия // Известия ЮФУ. Технические науки. – 2012. – № 12 (137). – С. 104-110.
3. *Бешта А.А., Кирпо М.А.* Построение модели доверия к объектам автоматизированной информационной системы для предотвращения деструктивных воздействий на систему // Известия Томского политехнического университета. Управление, вычислительная техника и информатика. – 2013. – Т. 322, № 5. – С. 104-108.
4. *Бешта А.А., Новикова Ю.В.* Способ численной оценки состояния автоматизированной информационной системы // Научно-технический вестник Поволжья. – 2013. – № 2. – С. 89-92.
5. *Бешта А.А.* Архитектура программного комплекса контроля над внутренним злоумышленником // Известия ЮФУ. Технические науки. – 2013. – № 12 (149). – С. 157-163.

Статью рекомендовал к опубликованию д.т.н., профессор О.Б. Макаревич.

Бешта Александр Александрович – Волгоградский государственный университет; e-mail: abewta@rambler.ru; 400062, г. Волгоград, пр-т Университетский, 100; тел.: 88442460368; кафедра информационной безопасности; старший преподаватель.

Цыбулин Анатолий Михайлович – e-mail: anatsybulin@yandex.ru; кафедра информационной безопасности; зав. кафедрой.

Beshta Alexander Alexandrovich – Volgograd State University; e-mail: abewta@rambler.ru; 100, Ave University, Volgograd, 400062, Russia; phone: +78442460368; the department of information security; lecturer.

Tsybulin Anatoly Mihaylovich – e-mail: anatsybulin@yandex.ru; the department of information security; head of department.

УДК 681.5:004(07)

В.И. Васильев, И.В. Шарабыров

ОБНАРУЖЕНИЕ АТАК В ЛОКАЛЬНЫХ БЕСПРОВОДНЫХ СЕТЯХ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Беспроводные сети передачи данных, в том числе и локального типа, продолжают стремительно развиваться. При этом зачастую безопасность в данных сетях не соответствует необходимому уровню. Одним из наиболее актуальных средств защиты от беспроводных атак являются системы обнаружения вторжений. 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. Приведен обзор сетевых атак, актуальных для локальных беспроводных сетей, а также методов интеллектуального анализа данных, которые можно использовать для обнаружения рассмотренных типов атак. В качестве методов интеллектуального анализа данных рассмотрены метод опорных векторов, метод k-ближайших соседей, нейронные сети и деревья принятия решений. Результаты экспериментов позволяют сделать вывод о практической значимости предложенного подхода к обнаружению атак в локальных беспроводных сетях.

Беспроводная сеть; сетевая атака; модель обнаружения; сигнатура; Wi-Fi.

V.I. Vasilyev, I.V. Sharabyrov

LOCAL WIRELESS NETWORKS ATTACKS DETECTION BASED ON INTELLIGENT DATA ANALYSIS

Nowadays wireless networks, including local ones, continue to evolve rapidly. Herewith security in these networks does not often correspond to the required level. One of the most actual protection means from the wireless attacks are the intrusion detection systems. Due to the extensive spread and wide possibilities of modern 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, as well as a comparison of data mining techniques, which can be used to detect the types of attacks mentioned above. The data mining techniques considered are support vector machine, k-nearest neighbor method, neural networks and decision trees. The experimental results allow making the conclusion about practical relevance of proposed approach for intrusion detection in local wireless networks.

Wireless network; network attack; detection model; signature; Wi-Fi.

Введение. Беспроводные сети завоевали огромную популярность. Их повсеместное распространение объясняется неоспоримыми преимуществами перед традиционными кабельными сетями: простота развертывания, мобильность пользователей в зоне действия сети, простое подключение новых пользователей. С другой стороны, безопасность таких сетей зачастую ограничивает их применение. 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 этих сетях:

- распространение сигнала за пределы контролируемой зоны;
- использование уязвимых протоколов и методов аутентификации;
- выпускаемые дополнения к стандартам не обеспечивают полноценную защиту от атак (например, протокол 802.11w не распространяется на контрольные кадры);
- ошибки в настройке различных компонентов беспроводной сети.

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

Для решения задачи обеспечения полноценной защиты многие компании активно ведут исследования в рамках беспроводных систем обнаружения вторжений (Wireless Intrusion Detection Systems, WIDS). Однако в данной области 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 современных методов интеллектуального анализа данных (ИАД) [3].

В [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.

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, включая использование методов ИАД в качестве основы для обнаружения данных атак.

1. Атаки, реализуемые в беспроводных сетях. 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.

В зависимости от цели атаки, характерные для семейства протоколов 802.11, можно поделить на несколько категорий [7]:

- получение несанкционированного доступа к сети: War Driving; ложные точка доступа или клиент (Rogue Access Point); подделка MAC-адреса (MAC Spoofing); взлом клиента беспроводной сети;
- нарушение целостности: инъекция поддельных кадров (802.11 Frame Injection); повтор, удаление пакетов с данными (802.11 Data Replay, Deletion); перехват и воспроизведение пакетов EAP, RADIUS (802.1X EAP Replay, 802.1X RADIUS Replay);
- нарушение конфиденциальности: подслушивание (Eavesdropping); атака «злой двойник» (Evil Twin); фишинг с помощью ложной точки доступа (AP Phishing); атака «человек посередине» (Man in the Middle);
- нарушение доступности: кража точки доступа; радиочастотное зашумление; захват среды ложными RTS/CTS-кадрами (Queensland DoS); наводнение запросами на подключение (Probe Request Flood); наводнение кадрами ассоциации, аутентификации, диссоциации, деаутентификации (Associate / Authenticate/ Disassociate/ Deauthenticate Flood); наводнение кадрами EAP (802.1X EAPStart, EAPFailure Flood);
- похищение данных аутентификации: взлом Pre-Shared Key; кража учетных данных 802.1X (802.1X Identity Theft); понижение уровня безопасности EAP (802.1X EAP Downgrade); взлом WPS PIN.

Данные атаки основаны на эксплуатации уязвимостей беспроводных сетей, представленных в базе WVE [8].

В качестве образцов атак транспортного и прикладного уровней можно воспользоваться усовершенствованной базой сигнатур NSL KDD-2009 [9], построенной на основе базы KDD-99 по инициативе американской Ассоциации перспективных оборонных научных исследований DARPA [10]. Для проведения исследований в области обнаружения вторжений был собран набор данных о соединениях, который охватывает широкий спектр различных вторжений, смоделированных в среде, имитирующей сеть Военно-воздушных сил США.

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 из четырех категорий атак: 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). Подробно атаки описаны в [11]. Соотношение числа атак разных типов показано в табл. 1, 2.

Таблица 1

Соотношение количества сигнатур атак в обучающей базе

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	40
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

Таблица 2

Соотношение количества сигнатур атак в тестовой базе

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

Некоторые из данных 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 вторжений (WIDS).

2. Системы обнаружения вторжений. В отличие от традиционных систем обнаружения вторжений, получающих все пакеты сети, беспроводные системы производят выборку сетевого трафика. Стандарты семейства 802.11 используют два основных диапазона частот: 2,4 и 5 ГГц, которые, в свою очередь, делятся на каналы. WIDS обеспечивают поочередное сканирование каналов на предмет наличия активных атак.

Схема функционирования WIDS представлена на рис. 1.

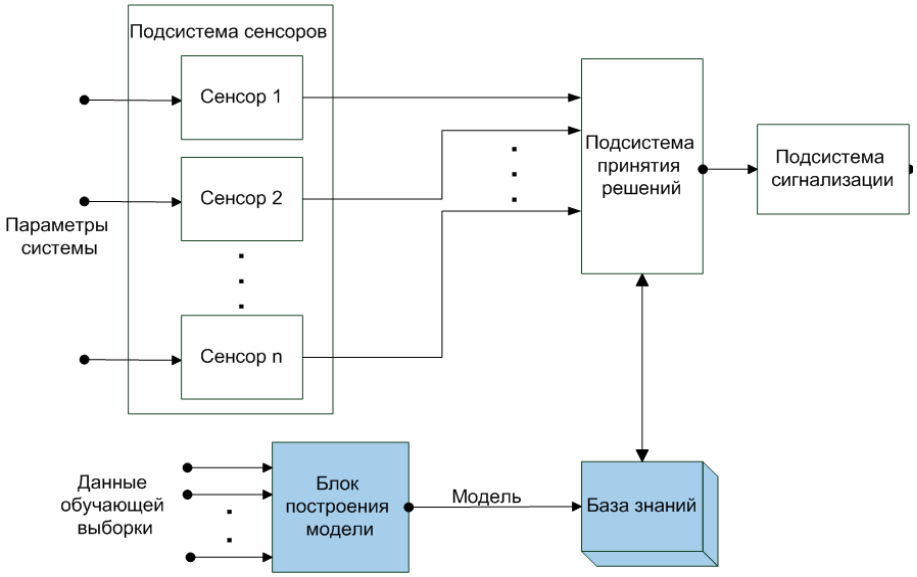

Рис. 1. Схема *функционирования WIDS*

Система работает в двух режимах:

- режим конфигурирования, когда в качестве входных данных в блок построения классифицирующей модели загружается набор сигнатур, представляющих собой пару {вектор параметров трафика | тип атаки}.
- режим нормальной работы, когда значения параметров трафика подаются в качестве входных данных на подсистему сенсоров. Далее подсистема принятия решений с помощью построенной на предыдущем этапе классифицирующей модели определяет, соответствуют ли показания сенсоров нормальному состоянию либо той или иной атаке, и, в зависимости от результата, срабатывает подсистема сигнализации.

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 принятия решения о безопасности какой-либо сетевой активности. В коммерческих продуктах это 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 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.

3. Методы интеллектуального анализа данных. В данной работе для выявления наиболее эффективного метода построения классифицирующей модели применительно к беспроводной системе обнаружения атак будет проведено сравнение четырех методов ИАД: метод опорных векторов, k-ближайших соседей, нейронные сети, деревья принятия решений.

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.

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 [12, 13] 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 распространённым 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.

4. Результаты экспериментов. Проектируемая WIDS структурно состоит из двух модулей:

- модуль обнаружения атак транспортного и прикладного уровней;
- модуль обнаружения атак канального и сетевого уровней.

Первый модуль на этапе обучения в качестве входных данных использует сигнатуры базы NSL KDD-2009. Для формирования обучающей выборки беспроводных атак канального и сетевого уровней была организована тестовая локальная беспроводная сеть с технологией защиты доступа WPA2-PSK. Для перехвата и

анализа пакетов использовался беспроводной адаптер Atheros AR9285 в режиме прослушивания. Собранные пакеты были проанализированы и приведены к виду, используемому в базе 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 неактуальности [14], в связи с чем количество признаков было сокращено. Выбранные признаки представлены в табл. 3, 4. 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 3

Значимые параметры трафика для первого модуля

Характеристика	Описание	Тип
Характеристики TCP-соединения		
duration	Продолжительность соединения (секунды)	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 правами администратора	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''	
rerror_rate / dst_host_same_src_port_rate	% соединений с ошибкой ``REJ'' / % соединений с одинаковым исходным портом	
same_srv_rate / dst_host_same_srv_rate	% соединений с одинаковым сервисом	
diff_srv_rate / dst_host_diff_srv_rate	% соединений с различным сервисом	
srv_count / dst_host_srv_count	Число соединений с совпадающим сервисом	
srv_serror_rate / dst_host_srv_serror_rate	% соединений с ошибкой ``SYN''	
srv_rerror_rate / dst_host_srv_rerror_rate	% соединений с ошибкой ``REJ''	
srv_diff_host_rate / dst_host_srv_diff_host_rate	% соединений с различающимися хостами	

Таблица 4

Значимые параметры трафика для второго модуля

Характеристика	Описание	Тип
Характеристики протоколов 802.11		
frame_type	Тип кадра	текстовый
SSID_empty	Длина тега SSID нулевая, 0 иначе	бинарный
invalid_chan_number	Неверный номер канала, 0 иначе	бинарный
more_fragments	Еще фрагменты для передачи, 0 иначе	бинарный
duration	Продолжительность передачи	численный
fragment_number	Номер фрагмента	численный
sequence_number	Номер кадра	численный
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

Эксперименты проводились в среде RapidMiner версии 5.3.015 [15] по схеме, приведенной на рис. 2.

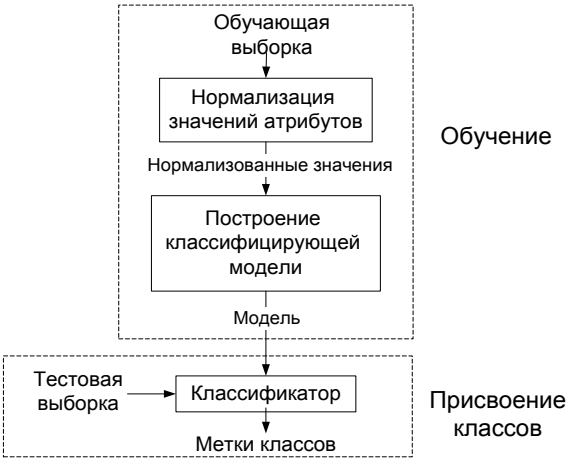

Рис. 2. Схема проведения эксперимента

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

После этого данные обучающей выборки поступали на вход блока построения соответствующей модели, на основе которой затем происходила классификация записей тестовой базы. На выходе классификатора формировалась тестовая выборка с предполагаемыми и действительными классами, на основании совпадения которых рассчитывалась точность и полнота обнаружения атак. Результаты классификации с помощью методов ИАД приведены в табл. 5.

Таблица 5

Показатели эффективности определения атак по классам в %

Класс	Метод опорных векторов		k-ближайших соседей		Нейронная сеть		Дерево принятия решений	
	Полнота	Точность	Полнота	Точность	Полнота	Точность	Полнота	Точность
neptune	98.99	99.98	99.81	99.38	99.66	99.78	99.76	97.83
normal	96.56	92.28	96.12	91.77	95.84	89.30	97.28	78.38
guess_passwd	77.25	100.00	65.39	100.00	78.72	78.27	0.32	100.00
smurf	100.00	99.70	99.10	100.00	95.34	98.91	100.00	88.55
satan	93.88	76.50	92.24	71.67	62.72	79.76	83.67	77.65
buffer_overflow	25.00	62.50	45.00	90.00	0.00	0.00	0.00	0.00
back	98.05	98.60	92.48	99.10	96.10	96.91	97.77	100.00
warezmaster	58.47	99.10	66.00	98.89	31.25	96.09	0.11	8.33
pod	95.12	72.22	95.12	72.22	80.49	68.75	75.61	77.50
nmap	98.63	94.74	100.00	93.59	100.00	92.41	98.63	100.00
ipsweep	97.16	93.84	98.58	95.86	97.87	89.03	98.58	83.73
portsweep	91.08	56.97	91.08	78.14	85.99	67.84	92.36	57.69
teardrop	83.33	21.28	83.33	21.28	83.33	21.28	100.00	24.49
land	57.14	100.00	14.29	100.00	57.14	100.00	71.43	100.00
Средняя		93.68		93.08		90.67		83.69

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 равна 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 одним скрытым слоем. Обучение продолжительностью 500 циклов производилось с помощью алгоритма обратного распространения ошибки. В качестве функции активации использовалась сигмоидальная функция. Величина максимальной ошибки обучения была ограничена значением 10^{-5}.

Построение дерева 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 – 21.

Как видно из табл. 5, методы опорных векторов и k-ближайших соседей показали близкие результаты в ходе обнаружения атак, несколько хуже проявили себя нейронная сеть и дерево принятия решений. Низкий процент обнаружения некоторых типов атак, таких как warezmaster, guess_passwd и buffer_overflow, 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 классифицирован неверно, поэтому результаты по ним не представлены в табл. 5.

Выводы. Представлен обзор сетевых атак, актуальных для локальных беспроводных сетей, рассмотрена схема функционирования беспроводной системы обнаружения вторжений, а также возможность применения методов ИАД для распознавания беспроводных атак.

В целом, рассмотренные методы ИАД показали высокую точность обнаружения в ходе проведения экспериментов, при этом наиболее точными оказались методы опорных векторов и k-ближайших соседей. Из этого 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 рассмотренных в данной статье методов ИАД.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

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).
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.
3. *Sinclair C., Pierce L., Matzner S.* An Application of Machine Learning to Network Intrusion Detection // Proceedings of Computer Security Applications Conference (ACSAC '99). – 1999. – P. 371-377.
4. *Tang H., Cao Z.* Machine Learning-based Intrusion Detection Algorithms // Journal of Computational Information Systems. – 2009. – P. 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).
6. *Mulay S., Devale P., Garje G.* Intrusion Detection System using Support Vector Machine and Decision Tree // International Journal of Computer Applications. – 2010. – Vol. 3, № 3. – P. 40-43.
7. *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/2cf7d7f61e47ae4ec 1257514004fce3f/$file/WLAN_Security% 20Risk%20Assessment%20and%20Countermeasures.pdf (дата обращения 12.03.2013).
8. WVE. Wireless Vulnerabilities and Exploits [Электронный ресурс]. URL: http://www.wve.org (дата обращения 05.10.2013).
9. The NSL-KDD Data Set. [Электронный ресурс]. URL: http://nsl.cs.unb.ca/NSL-KDD (дата обращения 22.01.2013).
10. KDD cup 99 Intrusion detection data set. [Электронный ресурс]. URL: http://kdd.ics.uci.edu/databases/kddcup99 (дата обращения 19.11.2011).
11. 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).
12. *Миронов К.В., Шарабыров И.В.* О применении метода опорных векторов в системах обнаружения атак // Мавлютовские чтения: Всероссийская молодежная научная конференция: сборник трудов в 5 т. Т. 3. – УГАТУ, 2012. – С. 28-30.
13. *Васильев В.И. [и др.].* Разработка модели обнаружения сигнатур атак на основе метода опорных векторов // Материалы XII Международной научно-практической конференции «Информационная безопасность-2012». Ч. 1. – Таганрог: Изд-во ТТИ ЮФУ, 2012. – С. 192-201.
14. *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, 2010. – Vol. 1. – P. 162-168.
15. RapidMiner Studio. [Электронный ресурс]. URL: https://rapidminer.com (дата обращения 01.09.2013).

Статью рекомендовал к опубликованию к.т.н., доцент А.А. Бакиров.

Васильев Владимир Иванович – Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Уфимский государственный авиационный технический университет»; e-mail: vasilyev@ugatu.ac.ru; 450000, г. Уфа, ул. К. Маркса, 12; тел.: 89173501139; кафедра вычислительной техники и защиты информации; зав. кафедрой; профессор.

Шарабыров Илья Викторович – e-mail: ilyashar@mail.ru; тел: 89196059521; кафедра вычислительной техники и защиты информации; аспирант.

Vasilyev Vladimir Ivanovich – Federal State Educational Institution of Higher Professional Education "Ufa State Aviation Technical University"; e-mail: vasilyev@ugatu.ac.ru; 12, K. Marx street, Ufa, 450000, Russia; phone: +79173501139; the department of computer engineering and information protection; head of the department; professor.

Sharabyrov Ilya Viktorovich – e-mail: ilyashar@mail.ru; phone: +79196059521; the department of computer engineering and information protection; postgraduate student.

УДК 004.056:061.68

В.М. Федоров, Д.П. Рублев

ИДЕНТИФИКАЦИЯ НАБИРАЕМОГО НА КЛАВИАТУРЕ ТЕКСТА ПО ВИБРОАКУСТИЧЕСКИМ ШУМАМ*

Рассмотрена проблема идентификация нажимаемых клавиш по виброакустическому сигналу, возникающему при работе пользователя с клавиатурой. Показаны возможность съёма идентифицирующих набранный текст данных по физическому (виброакустическому) каналу и преимущества данного метода, приведены модификации рабочего места оператора, описание стенда и программных средств. Рассмотрены особенности виброакустических сигналов, получаемых при работе пользователя с клавиатурой, произведён выбор устойчивых признаков, характеризующих нажатые клавиши из коэффициентов Фурье преобразования, кепстральных коэффициентов, коэффициентов линейного предсказания. Приведена схема синхронизации журнала клавиатурного регистратора и полученного виброакустического сигнала для получения точных сведений о локализации моментов нажатий/отпусканий клавиш и корректного формирования векторов признаков. Показана возможность идентификации клавиш по виброакустическим шумам, возникающим при наборе на клавиатуре, приведены зависимости оценок ошибок нейросетей от количества нейронов в скрытых слоях, выбора активационных функций и количества входных классов.

Виброакустический сигнал; дискретное Фурье преобразование; кепстральные коэффициенты; нейронные сети; коэффициенты линейного предсказания; идентификация.

V.M. Fedorov, D.P. Rublev

IDENTIFICATION OF TEXT TYPED ON KEYBOARD BY VIBROACOUSTICS NOISES

In this paper identification problem for pressed keys by vibroacoustic signal originated from user typing is reviewed. Capabilities of data collection for typed test restoration on vibroacoustic channel and advantages of this technique are shown. Operator's workplace modifications, stand description and software are considered. Features of vibroacoustic signals obtained from user's interaction with keyboard are reviewed, stable features which allow pressed key identification were selected from Fourier transform, cepstral and linear prediction coefficients. Sync scheme for keylogger's log file and vibroacoustics signal for precise localization of keypressing moments with key identification and feature vectors creation is reviewed. Keys identification by vibroacoustics typing noises is showed, dependencies of neural network errors on hidden layers neurons quantity, activation functions and output classes are considered.

Vibroacoustic signal; discrete Fourier transform; cepstral coefficients; neural networks; linear prediction coefficients; identification.

* Работа выполнена при поддержке гранта РФФИ № 12-07-00674-а.