УДК 004.4.056 Вестник УрФО № 2(28) / 2018, с. 43–48

Соколов А. Н., Алабугин С. К., Пятницкий И. А.

ПРИМЕНЕНИЕ МЕТОДОВ ОДНОКЛАССОВОЙ КЛАССИФИКАЦИИ ДЛЯ ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ

Современные тенденции в информационной безопасности характеризуются распространением комплексных целевых атак. Одним из способов борьбы со сложными сетевыми атаками, которые не всегда могут быть обнаружены классическими средствами защиты, является выявление аномалий в сетевом трафике. Для реализации такого подхода разумно использовать методы машинного обучения. В работе рассмотрены вопросы применения методов одноклассовой классификации как одной из техник машинного обучения для обнаружения вторжений, а также предложен вариант архитектуры системы обнаружения вторжения, основанной на таких методах.

Ключевые слова: *обнаружение вторжений, машинное обучение, обнаружение сетевых атак, одноклассовая классификация.*

Sokolov A. N., Alabugin S. K., Pyatnitsky I. A.

APPLYING OF ONE-CLASS CLASSIFICATION METHODS FOR INTRUSION DETECTION

The wider spread of complex target attacks is typical for the current state of information security. One way to deal with complex network attacks, which are difficult to detect by classic methods of protection, is the detection of anomalies in network traffic. It is reasonable to use machine learning methods, to implement this approach.

The paper deals with the application of one-class classification methods as one of the machine learning techniques for detecting intrusions, and proposes an architecture variant of an intrusion detection system based on such methods.

Keywords: *intrusion detection, machine learning, detection of network attacks, one-class classification.*

По данным компании Positive Technologies число компаний, столкнувшихся с целевыми атаками, за 2017 год увеличилось почти вдвое [1]. Кроме того, как отмечают эксперты, наблюдается рост сложности таких атак, в частности, злоумышленники активно применяют методы, затрудняющие анализ и расследование инцидентов.

В сложившейся ситуации классические методы, применяемые для обнаружения и защиты от вторжений [2], которые основаны на использовании сигнатур и правил, не всегда

в состоянии обнаружить атаку, не встречавшуюся ранее. Для обнаружения целевых атак предлагается использовать подход, основанный на выявлении аномалий сетевого трафика с применением техник и методов машинного обучения.

Существует два основных подхода, которые используются для обнаружения сетевых атак: обнаружение аномалий и сигнатурный подход [3]. В основе сигнатурного подхода лежит предположение, что любая сетевая атака может быть представлена уникальным шаблоном (сигнатурой атаки), а процесс обнаружения атак заключается в поиске их сигнатур. Этот подход обеспечивает малое число ложно положительных срабатываний (ошибок II рода): сигнатуры специально подбираются так, чтобы точно распознавать известные атаки. Однако в случае, если анализируется новая атака, для которой нет известной сигнатуры, её выявление проблематично. Кроме того, при реализации этого подхода, возникают такие сложности, как необходимость в постоянной актуализации базы известных атак и покрытия как можно большего числа родственных атак меньшим числом правил [3].

Обнаружение вторжений с помощью выявления аномалий предполагает построение систематически обновляющегося профиля нормальной активности характерной для сети, который содержит информацию о характерных сетевых пакетах и соединениях. Для построения профиля нормальной активности используются некоторые признаки (атрибуты) сетевого трафика [3]. При существенном расхождении наблюдаемых признаков с текущим профилем, делается вывод, что имеет место некая аномальная активность.

В основе подхода, основанного на обнаружении аномалий, лежит гипотеза, что к их появлению приводит любое вторжение. При этом аномалия может не быть атакой или вторжением. Появление аномалии также может быть связано с добавлением новых программных и аппаратных средств, неисправной работой сетевых устройств или деятельностью привилегированных пользователей сети. Поэтому описанный подход приводит к увеличению ошибок II рода. При этом если атака не приводит к появлению аномалий, она не может быть обнаружена. Кроме того, построение и систематическое обновление профиля нормальной активности, как правило, является трудоемкой и требовательной к вычислительным ресурсам задачей. Однако, применение подхода, основанного на выявлении аномалий, позволяет обнаруживать неизвестные сетевые атаки.

Таким образом, оба подхода имеют свои недостатки. Поэтому для обеспечения безопасности на практике разумно использовать их комбинацию.

Существует большое количество различных методов обнаружения сетевых атак, основанных на выявлении аномалий сетевого трафика [4]. Основной предпосылкой для применения методов машинного обучения в рамках этой задачи является отсутствие необходимости вручную создавать правила, согласно которым различаются аномалии и нормальная активность.

Под методами машинного обучения понимают, в частности, алгоритмы классификации (обучение с учителем) и кластеризации (обучение без учителя). Большинство работ в области обнаружения вторжений с применением машинного обучения используют алгоритмы классификации. Классические алгоритмы классификации, такие как метрические и линейные классификаторы, сети Байеса, деревья принятия решения и др. разработаны таким образом, чтобы классифицировать данные на несколько классов (например, классы «нормальный трафик» и «аномалия»). При этом обучаться такие алгоритмы также должны на данных, соответствующих разным классам.

Однако классические алгоритмы классификации по сути своей не приспособлены для обнаружения новых атак. Так, алгоритмы, обученные на выборке, состоящей из объектов, представляющих собой нормальный трафик и какие-либо виды сетевых атак, может классифицировать объект, представляющий новую для него атаку, как нормальный трафик. Кроме того, так как каждой компьютерной сети соответствует свой профиль нормального трафика, а алгоритм классификации каждый раз должен обучаться под конкретную компьютерную сеть, возникает проблема: как и где получить данные, соответствующие аномалиям (сетевым атакам) в конкретной компьютерной сети? Это решается, в частности, посредством генерации синтетических примеров данных, соответствующих аномалиям [3] или обучения с переносом опыта (transfer learning) [5].

В качестве альтернативного варианта, предлагается использовать алгоритмы одно-

классовой классификации. Одноклассовые классификаторы обучаются на примерах только одного класса, который в нашем случае соответствует нормальной активности, и строят решающее правило, согласно которому принимается решение о принадлежности объекта к нормальному классу [5].

К системе обнаружения вторжений (СОВ), использующей методы выявления аномалий, предъявляется ряд требований [3]:

1. Способность с высокой точностью выявлять аномалии в сети.
2. Малое количество ложных тревог.
3. Число входных параметров алгоритма должно быть небольшим, а их влияние на работу системы – низким.
4. Способность выявлять скрытые атаки.
5. Способность выявлять неизвестные системе атаки.

Процесс применения *СОВ на основе выявления аномалий характеризуется* четырьмя этапами [3]:

1. Сбор сетевого трафика.
2. Извлечение из трафика данных, представление их в виде признакового описания.
3. Анализ полученных данных и их разделение на нормальный и аномальный классы.
4. Обучение алгоритма на размеченных данных.

С учетом выше изложенных принципов и требований, предложена архитектура СОВ, использующей методы одноклассовой классификации. Схема ресурсов СОВ представлена на рис. 1.

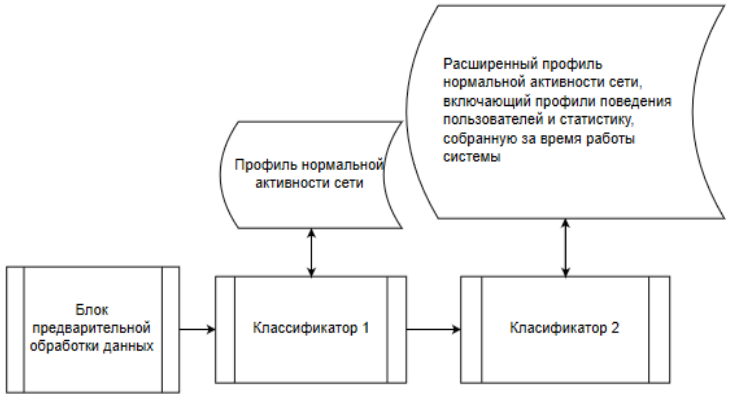

Рис. 1. Схема ресурсов системы обнаружения вторжений.

Отличительной особенностью предложенной системы является использование двух одноклассовых классификаторов: первый отвечает за распознавание типового нормального трафика, в случае же если он посчитает объект аномалией, решение будет приниматься вторым классификатором. Второй классификатор более сложный, медленный и работает с более большим количеством информации, включая информацию о состоянии сети и типичном поведении пользователей. Использование двух классификаторов в предложенной схеме мотивируется стремлением к одновременному обеспечению быстрой работы системы и высокой точности выявления аномалий. Для достижения этих целей первый классификатор должен отвечать следующим требованиям:

1. Высокая скорость обучения и работы.
2. Малый процент ложно отрицательных срабатываний (ошибок I рода).

Выберем алгоритм классификации, соответствующий сформулированным требованиям. Большинство алгоритмов одноклассовой классификации относятся либо к метрическим алгоритмам, которые основаны на вычислении функции расстояния в пространстве объектов, либо являются одной из большого числа модификаций метода опорных векторов. В ходе работы был рассмотрен ряд алгоритмов, обучающихся на примерах преимущественно одного класса. Чтобы выяснить, какой из них лучше подходит для использования в предложенной схеме СОВ, проведены эксперименты на наборах данных KDD'99 и NSL KDD [8], а также использованы результаты работ [9, 10]. Наборы данных KDD'99 и NSL KDD применяются для тестирования алгоритмов детектирования сетевых атак и широко используются исследователями.

В наборе NSL KDD содержатся: обучающая выборка с информацией о 125973 сетевых соединениях и тестовая выборка с информацией о 22544 сетевых соединениях.

Параметры, характеризующие каждое соединение, делятся на 4 группы:

1. Основные параметры каждого сетевого соединения.

2. Параметры, связанные с контентом каждого сетевого соединения.

3. Параметры, связанные с временными характеристиками каждого сетевого соединения.

4. Параметры, связанные с характеристиками хоста каждого сетевого соединения.

Полученные результаты представлены в табл. 1.

Таблица 1

Результатов реализации алгоритмов одноклассовой классификации на наборах данных KDD'99 и NSL KDD

Алгоритм	Набор данных	Точность	Доля неправильно распознанных атак
K-NN	KDD'99	0,974	0,72
Kth-NN	KDD'99	0,979	0,64
LOF	KDD'99	0,596	0,56
LOF-UB	KDD'99	0,577	0,55
COF	KDD'99	0,554	0,79
LoOP	KDD'99	0,574	0,78
INFLO	KDD'99	0,552	0,73
aLOCI	KDD'99	0,655	0,67
oc-SVM	KDD'99	0,951	0,26
η-oc-SVM	KDD'99	0,794	0,34
SVDD	NSL KDD	0,963	0,23
OCSVM	NSL KDD	0,941	0,24

По данным табл. 1 можно сделать вывод, что алгоритмы, основанные на методе опорных векторов (SVDD, OCSVM, oc-SVM, η-oc-SVM) имеют более высокую точность и распознают большее число атак. Однако, как и практически все модификации метода опорных векторов, эти алгоритмы сталкиваются с необходимостью во время обучения решать задачу квадратичного программирования [6], что негативно сказывается на их производительности. Кроме того, метод опорных векторов как модель алгоритма содержит большое количество параметров, которые необходимо варьировать в процессе обучения для достижения наилучшего результата. Это также добавляет сложностей при их практическом применении.

В качестве Классификатора 1 предложенной выше схемы предлагается использовать алгоритм, основанный на вычислении расстояния Махаланобиса [6, 7]. Он отчасти опирается на идеи, используемые в методе опорных векторов: используются понятие опорного вектора и отображение объектов в признаковое пространство большей размерности. При этом он спроектирован так, что решать задачу квадратичного программирования на стадии обучения не приходится.

В экспериментах с набором данных NSL KDD реализация описанного выше алгоритма на языке Python показала точность 0,95. Доля неправильно распознанных атак составила 0,102.

Таким образом, при анализе применимости методов одноклассовой классификации для обнаружения вторжений предложена схема работы системы обнаружения вторжений и исследованы возможности применения отдельных алгоритмов для выявления аномалий сетевого трафика на примере набора данных NSL KDD. Подводя итог, можно заключить, что применение методов одноклассовой классификации является перспективным способом повышения защищенности информационной системы от нетиповых, в том чисел целевых, атак.

Статья выполнена при поддержке Правительства РФ (Постановление №211 от 16.03.2013 г.), соглашение № 02.A03.21.0011.

Литература

1. Итоги года и прогнозы от Positive Technologies [Электронный ресурс]. – Режим доступа: https://www.ptsecurity.com/ru-ru/about/news/288913/, свободный (дата обращения 13.06.2018).

2. Мазиков К.И. Анализ современных сертифицированных средств обнаружения вторжений в информационных сетях // Вестник Тамбовского университета. Серия: Естественные и технические науки. – 2014. – №2. – С. 661-662.

3. Bhuyan M. H., Bhattacharyya D. K., Kalita J. K. Network Traffic Anomaly Detection: Concepts, Techniques, and Tools, Springer, 2017, 262 p.

4. Браницкий А. А., Котенко И. В. Анализ и классификация методов обнаружения сетевых атак, Тр. СПИИРАН, 45 (2016), С. 207–244.

5. Zhao, J., Shetty, S. and Pan, J.W., 2017, October. Feature-based transfer learning for network security. In Military Communications Conference (MILCOM), MILCOM 2017-2017 IEEE (pp. 17-22). IEEE..

6. Nader P., Honeine P., Beauseroy P. Mahalanobis-based one-class classification //Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop on. – IEEE, 2014. – С. 1-6.

7. Nader P., Honeine P., Beauseroy P. Online one-class classification for intrusion detection based on the mahalanobis distance //Proc. 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. – 2015. – С. 1-6.

8. Tavallaee M. et al. A detailed analysis of the KDD CUP 99 data set //Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on. – IEEE, 2009. – С. 1-6.

9. Goldstein M., Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data //PloS one. – 2016. – Т. 11. – №. 4. – С. 152-173.

10. Kumar S., Nandi S., Biswas S. Research and application of one-class small hypersphere Support Vector Machine for Network anomaly detection //Communication Systems and Networks (COMSNETS), 2011 Third International Conference on. – IEEE, 2011. – С. 1-4.

Refereces

1. Itogi goda i prognozy ot Positive Technologies. Available at: https://www.ptsecurity.com/ru-ru/about/news/288913/ (acessed 13.06.2018).

2. Mazikov K.I. Analiz sovremennykh sertifitsirovannykh sredstv obnaruzheniya vtorzheniy v informatsionnykh setyakh // Vestnik Tambovskogo universiteta. Seriya: Estestvennyye i tekhnicheskiye nauki. – 2014. – №2. – P. 661-662.

3. Bhuyan M. H., Bhattacharyya D. K., Kalita J. K. Network Traffic Anomaly Detection: Concepts, Techniques, and Tools, Springer, 2017, 262 p.

4. Branitskiy A. A., Kotenko I. V. Analiz i klassifikatsiya metodov obnaruzheniya setevykh atak, Tr. SPIIRAN, 45 (2016). P. 207–244.

5. Zhao, J., Shetty, S. and Pan, J.W., 2017, October. Feature-based transfer learning for network security. In Military Communications Conference (MILCOM), MILCOM 2017-2017 IEEE (pp. 17-22). IEEE.

6. Nader P., Honeine P., Beauseroy P. Mahalanobis-based one-class classification //Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop on. – IEEE, 2014. – P. 1-6.

7. Nader P., Honeine P., Beauseroy P. Online one-class classification for intrusion detection based on the mahalanobis distance //Proc. 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. – 2015. – P. 1-6.

8. Tavallaee M. et al. A detailed analysis of the KDD CUP 99 data set //Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on. – IEEE, 2009. – P. 1-6.

9. Goldstein M., Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data //PloS one. – 2016. – T. 11. – №. 4. – P. 152-173.

10. Kumar S., Nandi S., Biswas S. Research and application of one-class small hypersphere Support Vector Machine for Network anomaly detection //Communication Systems and Networks (COMSNETS), 2011 Third International Conference on. – IEEE, 2011. – P. 1-4.

СОКОЛОВ Александр Николаевич, кандидат технических наук, доцент, заведующий кафедрой защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д 76. E-mail: sokolovan@susu.ru

АЛАБУГИН Сергей Константинович, аспирант кафедры защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д 76. E-mail: sergei_alabugin@mail.ru

ПЯТНИЦКИЙ Илья Альбертович, аспирант кафедры защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д 76. E-mail: Ankidoom@gmail.com

SOKOLOV Alexander, Candidate of Engineering Science, Docent, Head of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: sokolovan@susu.ru

ALABUGIN Sergei, postgraduate student of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: sergei_alabugin@mail.ru

PYATNITSKY Ilya, postgraduate student of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: Ankidoom@gmail.com