УДК 004.056 DOI: 10.18698/2541-8009-2021-03-684

АНСАМБЛЬ КЛАССИФИКАТОРОВ ДЛЯ СИСТЕМЫ ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ

А.А. Ломанов

a.a.lomanov@gmail.com
SPIN-код: 8959-1032

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Рассмотрены различные типы сетевых атак и методы их обнаружения. Проведено сравнение различных методов классификации сетевых атак и выбраны классификаторы для ансамбля. Основное внимание уделено описанию предлагаемого решения и структуры ансамбля классификаторов для системы обнаружения вторжений. Рассмотрены наиболее популярные наборы данных для тестирования. Для обучения и тестирования выбран набор данных CICIDS2017. Тестирование показало преимущество применения разработанного подхода перед использованием одиночного классификатора в точности распознавания сетевой атаки, что подтверждает эффективность его применения в системе обнаружения вторжений для обнаружения и классификации сетевой атаки.

Ключевые слова

Сетевые атаки, сигнатурный метод, обнаружение вторжений, классификация, терминальный классификатор, гибридизация, ансамбль классификаторов, машинное обучение

Поступила в редакцию 10.02.2021

Введение. В настоящее время функционирование любой организации или предприятия практически полностью зависит от стабильной работы сетевого оборудования и устройств. При этом к такой сети предъявляют требования по безопасности передачи информации и доступа к ней, поскольку информационная безопасность — ключевая составляющая уникальности и конкурентоспособности организации. В связи с этим возникла актуальная проблема обнаружения отклонений в работе сетевых устройств. Существующие решения имеют ряд недостатков, которые не позволяют с наибольшей эффективностью обнаруживать вторжения в сети [1]. По этой причине актуальной является задача нахождения более эффективного решения обнаружения вторжений.

В статье предложен новый подход к решению этой задачи, который основан на использовании гибридной системы, сочетающей в себе систему обнаружения вторжений, систему защиты конечного узла и ансамбля классификаторов. В качестве классификаторов были выбраны деревья решений, метод k-ближайших соседей и наивный байесовский классификатор, что обусловлено невысокой сложностью их обучения, неплохой точностью классификации и относительной простотой реализации.

Проведенное тестирование системы показало достаточно неплохую точность (более 70 % корректных распознаваний сетевых атак) каждого классифи-

катора и значительную точность (более 90 % корректных распознаваний сетевых атак) при использовании ансамбля.

В данной статье рассмотрены существующие работы и наборы данных, опыт использования которых применен для создания предлагаемого подхода, далее следуют описание и обоснование предлагаемого подхода, затем описано проведенное тестирование и сделано заключение об эффективности предложенного подхода.

Классификация методов обнаружения сетевых атак. Классической классификацией методов обнаружения сетевых атак является интерпретация по входным данным, а именно разделение на методы обнаружения злоупотреблений и методы обнаружения аномалий. В первом случае подразумевается обнаружение определенной известной сигнатуры путем мониторинга сетевого трафика. Второй случай подразумевает моделирование нормального поведения сети, и отклонения от этой модели считаются аномальными.

Общая схема обнаружения сетевых атак в описанной классификации представлена на рис. 1 [2].

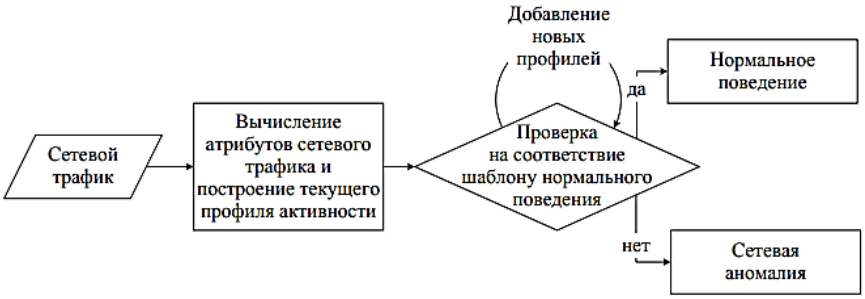

Рис. 1. Общая схема обнаружения сетевых атак

Общий алгоритм выявления сетевых аномалий можно описать следующим образом. Данными для анализа является сетевой трафик, представленный как набор сетевых пакетов, в общем случае фрагментированных на уровне IP. Собранные сырые данные в дальнейшем послужат источником при формировании необходимой информации для последующего анализа. Далее производится нормализация данных за определенный период времени. Созданный набор данных рассматривается на предмет выявления отклонений.

Ансамбль классификаторов. Ансамбль структурно состоит из одного или более модулей анализа — классификаторов. Наличие нескольких классификаторов требуется для повышения эффективности обнаружения. Каждый классификатор выполняет поиск атак или вторжений определенного типа. Входными данными для классификатора является информация из подсистемы сбора информации или от другого классификатора. Результат работы подсистемы — индикация о состоянии защищаемой системы. Когда классификатор сообщает об обнаружении несанкционированных действий, на его выходе может появляться некоторая дополнительная информация. Обычно эта информация содержит выводы, подтверждающие факт наличия вторжения или атаки.

Существует несколько методов построения ансамбля классификаторов [3]. Наиболее популярными являются методы “bagging” и “boosting” [4], которые основаны на манипуляциях с исходным обучающим множеством в целях построения нескольких классификаторов. Для построения независимых классификаторов наиболее эффективным методом является обучение отдельных членов ансамбля на различающихся подмножествах признаков. Таким образом, построение ансамбля классификаторов на основе декомпозиции исходного набора признаков, описывающих объекты данных, в большинстве случаев имеет преимущества.

Однако когда размерность признакового пространства достаточно большая, такой способ является неэффективным. В связи с этим существует также подход к построению ансамбля классификаторов, отличительной особенностью которого является использование генетического алгоритма для одновременного отбора нескольких подмножеств признаков для построения отдельных классификаторов, входящих в состав ансамбля. Использование генетического алгоритма при решении оптимизационной задачи декомпозиции исходного множества признаков для построения ансамбля классификаторов объясняется простотой кодирования решения оптимизационной задачи, отсутствием ограничений на гладкость оптимизируемой функции (что позволяет в качестве последней использовать точность классификации с использованием ансамбля) и отсутствием эффективных субоптимальных алгоритмов.

Описание решения. Несмотря на все достоинства систем обнаружения вторжений (Intrusion Detection System — IDS [5]), все существующие системы имеют один существенный недостаток. В крупных корпоративных сетях приходится постоянно дорабатывать правила и настройки IDS-системы и приводить их в соответствие с изменяющейся конфигурацией сети. При этом крайне сложно в полной мере учесть специфику сети, что приводит к большому количеству ложных срабатываний. Таким образом, все достоинства систем обнаружения вторжений в крупных сетях не реализуются.

Для улучшения качества работы защитной системы предлагается использовать ансамбль классификаторов. Ансамбль классификаторов — это набор методов классификации активности в сети, результаты которых анализируются совместно [6]. Таким образом достигаются наибольшая точность и уверенность в определении аномальной активности.

Ансамбль состоит из определенного числа классификаторов и блока настройки правил. На вход каждого классификатора поступают одинаковые входные данные. В зависимости от функции, заданной для классификатора, его выходом может быть либо степень уверенности, либо булево значение «обнаружено» или «не обнаружено». Выходные значения классификаторов обрабатываются блоком правил ансамбля. Этот блок вычисляет значения степеней уверенности в принадлежности события к определенному классу. В соответствии с настройкой и в зависимости от вычисленных степеней уверенности блок правил решает, к какому классу с наибольшей точностью принадлежит рассмотренное событие.

Структура системы. Ансамбль классификаторов включает в себя три классификатора и блок настройки правил.

Для обнаружения известных аномалий разумно использовать сигнатурный метод обнаружения. При таком подходе надежно детектируются шаблоны сигнатур известных типов атак с помощью средств IDS. Однако для обнаружения аномалий, сигнатуры которых отсутствуют в базе шаблонов, требуется применять комбинированную систему обнаружения вторжений. Структурная схема такой системы представлена на рис. 2.

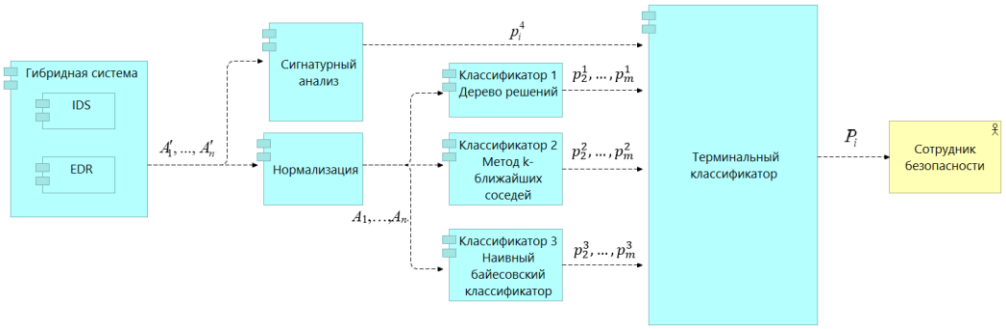

Рис. 2. Структурная схема комбинированной системы

На первом уровне данной структуры установлена гибридная система, состоящая из IDS и EDR (Endpoint Detection & Response — система обнаружения на конечной точке), т. е. системы обнаружения аномалий на конечном узле. Эти компоненты собирают пакеты данных сетевого трафика и состояния конечных устройств. Количество таких событий зависит от архитектуры конкретной сети, настройки систем для этой сети и ситуации в сети. Таким образом, мы имеем n событий $A'_1, \ldots, A'_n$. Затем события проходят процесс нормализации и, по возможности, снижения размерности. Такая обработка может помочь ускорить обнаружение аномалий и понизить сложность классификации. Для сигнатурного анализа, однако, обработка не требуется. События анализируются на предмет наличия в них сигнатуры известной аномалии, результаты p_i^4 поступают сразу на вход терминального классификатора.

На втором уровне данной структуры располагаются три классификатора. В качестве входа для них используются нормализованные параметры $A_1,\ldots,A_n$. Выходы классификаторов представляют собой набор дискретных величин $p_2^3, \ldots, p_m^3$ $p_2^k, \ldots, p_m^k$. Каждая величина представляет собой метку класса S_2, …, S_m. Метка представляет собой значение в диапазоне [0; 1]. Метка класса S_1 означает, что текущее состояние соответствует нормальному.

При выборе классификаторов учитывали сложность их обучения, точность классификации и простоту реализации. Было проведено сравнение таких алгоритмов, как дерево решений (DT), метод опорных векторов, метод k-ближайших соседей (KNN) и наивный байесовский классификатор (NB) [7]. Деревья решений имеют достаточно невысокую сложность обучения, достаточ-

но точны и просты в реализации. Метод *k*-ближайших соседей имеет относительно высокую вычислительную сложность, но его очень легко реализовать и интерпретировать, а также он имеет высокую точность, поэтому имеет смысл использовать этот метод. Наивный байесовский классификатор, как известно, максимально прост, достаточно точен и эффективен на наборе независимых переменных, а также не представляет вычислительной сложности, поэтому его использование объективно разумно. Метод опорных векторов [8] имеет высокую сложность обучения, при этом он достаточно сложен в реализации, поэтому не имеет смысла применять этот классификатор.

На третьем уровне данной структуры находится терминальный классификатор, который работает по принципу взвешенного голосования, а именно агрегирует выходы классификаторов второго уровня и относит текущее состояние к одному из классов. Взвешенное голосование основано на применении весовых коэффициентов. Выход терминального классификатора представляет собой множество значений P_m, каждое из которых отражает принадлежность текущего состояния S_m и определяется по формуле

$$P_i = \max\left(\operatorname{sign}[\alpha_1 p_i^1 + \alpha_2 p_i^2 + \alpha_3 p_i^3 0,\ p_i^4\right),$$

где P принимает значение в диапазоне [0; 1], $i = 2, \ldots, m$; α_1, α_2, α_3 — весовые коэффициенты, выбираемые на основе точности каждого классификатора.

Описание результатов тестирования. Существует несколько основных наборов данных для распознавания сетевых атак [9]. Наиболее популярными являются KDD99, NSL-KDD, CDX 2009. Набор KDD99, как понятно из его названия, является относительно устаревшим из-за срока давности. Также он имеет ряд недостатков, которые были устранены в наборе NSL-KDD, где уже устранены избыточность и дублирование данных, а также применен более продуманный подход к формированию выборок. Однако данный набор попрежнему не самый современный и актуальный. Набор ITOC CDX 2009 представляет собой набор логов для Snort IDS и Apache Web Server, что не совсем подходит для текущей задачи, необходим более универсальный набор. В связи с этим был выбран набор CICIDS2017 [10]. Данный набор является наиболее современным и актуальным из найденных, он также содержит наиболее разнообразные типы атак и позволяет провести наиболее обширный эксперимент.

Для оценки точности и эффективности классификаторов был проведен эксперимент. Работоспособность была проверена на наборе данных CICIDS2017. Указанный набор данных содержит периоды нормального состояния и самые современные распространенные атаки, которые напоминают настоящие реальные данные. Он также включает результаты анализа сетевого трафика с использованием python-анализатора CICFlowMeter с помеченными потоками на основе метки времени, IP-адресов источника и назначения, портов источника и назначения, протоколов и атак. Также доступно определение извлеченных функций.

Набор содержит данные о состоянии в период за 5 дней. Первый день характеризуется нормальным состоянием системы без каких бы то ни было аномалий. Данные следующих дней содержат сведения о таких типах атак, как Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet, DDoS и др. Результаты классификации по каждому типу атак для каждого классификатора приведены в таблице.

Точность распознавания классификаторами различных типов сетевых атак, %

	Normal	DDoS	Brute Force	Web Attack	Others
DT	77,55	86,29	75,13	88,31	66,39
KNN	79,13	87,48	77,25	90,18	67,84
NB	72,86	82,56	70,61	84,93	65,72
DT+KNN	81,32	91,33	79,37	91,78	69,12
DT+NB	79,47	87,51	75,88	90,85	67,51
KNN+NB	80.61	87,93	78,66	92,04	68,07
All	86,93	95,73	85,53	98,66	76,48

По результатам эксперимента можно сказать, что, несмотря на достаточно высокую сложность классификации данных в наборе CICIDS2017, каждый отдельный классификатор показал неплохую точность. Хуже всех показал себя наивный байесовский классификатор, что вполне объяснимо его простотой. Однако этот классификатор позволяет заметно повысить точность распознавания при использовании совместно с другими классификаторами, и в этом случае его простота становится преимуществом, поскольку не происходит значительного усложнения реализации комбинированных классификаторов. Важнейший результат эксперимента — доказательство того факта, что при использовании ансамбля классификаторов вместо одного классификатора можно добиться значительного повышения точности.

Заключение. В рамках данной работы проведено исследование различных подходов к обнаружению и классификации типов сетевых атак. Описано решение обнаружения и классификации сетевых атак на основе гибридной системы, включающей в себя IDS, EDR и ансамбль классификаторов. Выполнено сравнение различных классификаторов и обоснован выбор используемых в предложенной системе. Работоспособность системы проверена на наборе данных CICIDS2017, эксперимент подтвердил достаточно высокую точность классификации, что свидетельствует о возможности использования предложенной системы для обнаружения вторжений.

Литература

[1] Кеммерер Р., Виджна Д. Обнаружение вторжений: краткая история и обзор. *Открытые системы. СУБД*, 2002, № 7–8. URL: https://www.osp.ru/os/2002/07-08/181714

[2] Браницкий А.А., Котенко И.В. Анализ и классификация методов обнаружения сетевых атак. *Труды СПИИРАН*, 2016, № 45, с. 207–244. DOI: https://doi.org/10.15622/sp.45.13

[3] Новоселова Н.А., Том И.Э. Подход к построению ансамбля классификаторов с использованием генетического алгоритма. *Искусственный интеллект*, 2009, № 3. URL: http://dspace. nbuv.gov.ua/bitstream/handle/123456789/8021/09-Novoselova. pdf? sequence=1 (дата обращения: 26.09.2020).

[4] Pham N., Foo E., Suriadi S., et al. Improving performance of intrusion detection system using ensemble methods and feature selection. *Proc. ACSW Multiconf.*, 2018, art. 2. DOI: https://doi.org/10.1145/3167918.3167951

[5] Ahmad Z., Khan A.S., Shiang C.W., et al. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Trans. Emerg. Telecommun. Technol.*, 2021, vol. 32, no. 1, art. e4150. DOI: https://doi.org/10.1002/ett.4150

[6] Gao X. Shan C., Hu C., et al. An adaptive ensemble machine learning model for intrusion detection. *IEEE Access*, 2019, vol. 7, pp. 82512–82521. DOI: https://doi.org/10.1109/ ACCESS.2019.2923640

[7] Алгулиев Р., Алыгулиев Р., Имамвердиев Я. и др. Обнаружение DoS атак с применением ансамбля классификаторов. URL: https://ict.az/uploads/konfrans/info_sec/RS02_DOS-ATTACKS-DETECTION-USING-AN-ENSEMBLE-OF-CLASSIFIERS.pdf (дата обращения: 12.11.2020).

[8] Sahu S.K., Katiyar A., Kumari K.M., et al. An SVM-based ensemble approach for intrusion detection. *IJITWE*, 2019, vol. 14, no. 1, pp. 66–84. DOI: https://doi.org10.4018/IJITWE.2019010104

[9] Бурлаков М.Е. Применение метода анализа соответствий для оптимизации комбинаций атрибутов у наборов данных. *Вестник ПНИПУ*, 2018, № 26, с. 7–28.

[10] Intrusion detection evaluation dataset (CICIDS2017). *unb.ca: веб-сайт.* URL: https://www.unb.ca/ cic/datasets/ids-2017.html (дата обращения: 15.11.2020).

Ломанов Александр Александрович — студент кафедры «Информационные системы и телекоммуникации», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Сакулин Сергей Александрович — кандидат технических наук, доцент кафедры «Информационные системы и телекоммуникации», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Ссылку на эту статью просим оформлять следующим образом:
Ломанов А.А. Ансамбль классификаторов для системы обнаружения вторжений. *Политехнический молодежный журнал*, 2021, № 03(56). http://dx.doi.org/ 10.18698/ 2541-8009-2021-03-684

INTRUSION DETECTION SYSTEM CLASSIFIER ENSEMBLE

A.A. Lomanov a.a.lomanov@gmail.com
SPIN-code: 8959-1032

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

Various types of network attacks and methods of their detection are considered. Comparison of various methods of classification of network attacks is carried out and classifiers for the ensemble are selected. The main attention is paid to the description of the proposed solution and the structure of the classifiers ensemble for the intrusion detection system. The most popular data sets for testing are considered. The CICIDS2017 dataset was selected for training and testing. Testing has shown the advantage of using the developed approach over the use of a single classifier in the accuracy of recognizing a network attack, which confirms the effectiveness of its application in an intrusion detection system for detecting and classifying a network attack.

Keywords

Network attacks, signature method, intrusion detection, classification, terminal classifier, hybridization, ensemble of classifiers, machine learning

Received 10.02.2021

References

[1] Kemmerer R., Vidzhna D. Intrusion detection: short history end review. *Otkrytye sistemy. SUBD* [Open Systems. DBMS], 2002, no. 7–8. URL: https://www.osp.ru/os/2002/07-08/181714 (in Russ.).

[2] Branitskiy A.A., Kotenko I.V. Analysis and classification of methods for network attack detection. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2016, no. 45, pp. 207–244. DOI: https://doi.org/10.15622/sp.45.13 (in Russ.).

[3] Novoselova N.A., Tom I.E. Design of classifier ensemble by genetic algorithm (in Russ.) .*Iskusstvennyy intellect* [Artificial Intelligence], 2009, no. 3. URL: http://dspace.nbuv. gov.ua/bitstream/handle/123456789/8021/09-Novoselova.pdf?sequence=1 (accessed: 26.09.2020).

[4] Pham N., Foo E., Suriadi S., et al. Improving performance of intrusion detection system using ensemble methods and feature selection. *Proc. ACSW Multiconf.*, 2018, art. 2. DOI: https://doi.org/10.1145/3167918.3167951

[5] Ahmad Z., Khan A.S., Shiang C.W., et al. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Trans. Emerg. Telecommun. Technol.*, 2021, vol. 32, no. 1, art. e4150. DOI: https://doi.org/10.1002/ett.4150

[6] Gao X. Shan C., Hu C., et al. An adaptive ensemble machine learning model for intrusion detection. *IEEE Access*, 2019, vol. 7, pp. 82512–82521. DOI: https://doi.org/10.1109/ACCESS.2019.2923640

[7] Alguliev R., Alyguliev R., Imamverdiev Ya., et al. Obnaruzhenie DoS atak s primeneniem ansamblya klassifikatorov [DoS intrusion detection using suit of qualifiers] (in Russ.). URL: https://ict.az/uploads/konfrans/info_sec/ rs02_dos-attacks-detection-using-an-ensemble-of-classifiers.pdf (accessed: 12.11.2020).

[8] Sahu S.K., Katiyar A., Kumari K.M., et al. An SVM-based ensemble approach for intrusion detection. *IJITWE*, 2019, vol. 14, no. 1, pp. 66–84. DOI: https://doi.org10.4018/IJITWE.2019010104

[9] Burlakov M.E. Application the method of correspondence analysis to optimize combinations of attributes from datasets. *Vestnik PNIPU* [PNRPU Bulletin], 2018, no. 26, pp. 7–28 (in Russ.).
[10] Intrusion detection evaluation dataset (CICIDS2017). *unb.ca: website*. URL: https://www.unb.ca/ cic/datasets/ids-2017.html (accessed: 15.11.2020).

Lomanov A.A. — Student, Department of Information Systems and Telecommunications, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Sakulin S.A., Cand. Sc. (Eng.), Assoc. Professor, Department of Information Systems and Telecommunications, Bauman Moscow State Technical University, Moscow, Russian Federation.

Please cite this article in English as:
Lomanov A.A. Intrusion detection system classifier ensemble. *Politekhnicheskiy molodezhnyy zhurnal* [Politechnical student journal], 2021, no. 03(56). http://dx.doi.org/ 10.18698/2541-8009-2021-03-684.html (in Russ.).