ОРГАНИЗАЦИЯ НЕЙРОСЕТЕВОЙ АРХИТЕКТУРЫ АВТОКОДИРОВЩИКА ПРИ ВЫДЕЛЕНИИ ПРИЗНАКОВ DOS/DDOS-АТАК

Мостовщиков Д.Н.[1], Дос Е.В.[2], Камалиденов К.Ш.[3]

[1]Мостовщиков Дмитрий Николаевич - старший системный архитектор;
[2]Дос Евгений Владимирович - старший системный архитектор,
Li9, Inc.,
г. Феникс, Соединенные Штаты Америки;
[3]Камалиденов Куаныш Шарипханович - ведущий системный архитектор,
Digital IQ, г. Нур-Султан, Республика Казахстан

Аннотация: *рассмотрены актуальные подходы по организации системы киберзащиты информационных ресурсов сетевых сервисов. Указано, что современные сетевые сервисы базируются на распределенной архитектуре и динамической маршрутизации, что приводит к расширению набора уязвимостей, в частности, таких как неправомерный доступ к конфиденциальным данным, нарушение стабильной работы сервиса, а также внедрение вредоносного программного кода. Проведена классификация методов защиты от DoS- и DDoS-атак, а также определены особенности построения нейросетевых алгоритмов, которые используются с целью выделения признаков кибератак. Была предложена методика построения нейросетевой архитектуры на основе каскадного автокодировщика с глубинным обучением, отмечено, что соответствующие алгоритмы позволяют отслеживать признаки DoS- и DDoS-атак в режиме реального времени. Результаты исследования дали возможность разработать комплексную методологию для решения задач по построению системы защиты облачного сервиса на основе распределенной сетевой архитектуры при помощи нейросетевых алгоритмов отслеживания признаков кибератак.*

Ключевые слова: *сетевой сервис, информационный ресурс, нейросетевая архитектура, архитектура «автокодировщик», DoS/DDoS-атака, математическая модель, целевые показатели.*

ORGANIZATION OF THE CASCADE AUTOENCODER ANN ARCHITECTURE FOR TRACKING OF THE DOS/DDOS ATTACKS

Mostovshchikov D.N.[1], Dos E.V.[2], Kamalidenov K.Sh.[3]

[1] Mostovshchikov Dmitrii Nikolayevich – Senior Systems Architect;
[2] Dos Evgenii Vladimirovich - Senior Systems Architect,
LI9, INC.,
PHOENIX, AZ, UNITED STATES OF AMERICA;
[3]Kamalidenov Kuanysh Sharipkhanovich - Senior Systems Architect,
DIGITAL IQ, NUR-SULTAN, REPUBLIC OF KAZAKHSTAN

Abstract: *approaches to organizing a cyber-protection system for information resources of network services are analyzed. It is indicated that modern network services are based on a distributed architecture and dynamic routing, which leads to an expansion of the set of vulnerabilities, in particular, unauthorized access to confidential data, disruption of the stable operation of the service, and the use of malicious software. The classification of methods of protection against DoS- and DDoS-attacks, as well as the features of the construction of Artificial Neural Network (ANN) algorithms, which are distinguished by a high degree of identification of cyber-attacks patterns, has been carried out. A technique was proposed for constructing a neural network architecture based on a deep learning cascade autoencoder, algorithms based on which allow tracking DoS and DDoS attacks patterns in real time. Optimization of algorithms for detecting the cyber-attacks was carried out through leveling errors typical for the operation and training of ANNs. The results of the study made it possible to develop a comprehensive methodology for solving problems of the cloud service based on a distributed network architecture protection system organizing with application of neural network algorithms for tracking cyber-attacks patterns.*

Keywords: *network service, information resource, neural network architecture, stacked autoencoder, DoS/DDoS-attack, mathematical model, targets.*

УДК 004.07

Введение

Активное развитие сетевых сервисов [1-3] и, в частности, активное внедрение облачных сервисов, которое наблюдается последние два десятилетия, характеризуется использованием архитектуры распределенной информационной системы (Distributed Information System, DIS) и протоколов динамической маршрутизации (Dynamic Routing Protocol, DRP). Указанные тенденции, в свою очередь, привели к появлению и расширению разного рода уязвимостей, в частности таких групп классификации внешних угроз как:

- неправомерный доступ к конфиденциальным данным и данным протоколов системных служб сетевого сервиса через организацию скрытых каналов передачи данных [4, 5];
- нарушение стабильной работы сетевого сервиса через увеличение нагрузки на локальные вычислительные узлы, а также внесение изменений в программные алгоритмы и данные протоколов [6, 7];

- внедрение вредоносного программного кода, который на долгосрочной основе контролирует сетевой сервис в соответствии с задачами злоумышленников [8, 9].

Проведение исследования в соответствии с задачей организации эффективной системы защиты сетевого сервиса дало возможность указать на приоритет в построении алгоритмов выделения признаков DoS- и DDoS-атак (Denial of Service — атака типа «отказ в обслуживании» и Distributed Denial of Service — распределённая атака типа «отказ в обслуживании»). Данный вид атак является наиболее распространенным в связи с относительной простотой их реализации через генерацию большого количества запросов на один или несколько вычислительных узлов системы, имитирующих запросы пользователей сетевого сервиса. С целью отслеживания признаков DoS- и DDoS-атак широко используются нейросетевые алгоритмы с архитектурой типа каскадный автокодировщик (Stacked AutoEncoder, SAE) и глубинным нейронным обучением (Deep Neural Network, DNN), которые при соответствующей настройке эффективно отслеживают признаки данного вида кибератак, анализируя сетевой трафик в режиме реального времени [10, 11].

Анализ последних исследований и публикаций в области организации сетевых сервисов [1-3] и построения систем защиты соответствующих информационных ресурсов от внешних угроз [4-9] указывает на особенности использования нейросетевых алгоритмов при выделении признаков DoS- и DDoS-атак [10, 11, 15-16]. В частности были рассмотрены статистические данные исследований, в которых использовались системы глубинного обучения на этапах (i) сокращения набора признаков, (ii) оптимизации гиперпараметров (iii) классификации элементов входного набора [12-14]. При этом перспективными представляются методы обучения в соответствии со схемой «без учителя» с использование нейросетевой архитектуры типа автокодировщик. Это позволяет на стандартном обучающем наборе после проведения этапов предварительного обучения (формирование пространства признаков) и тонкой настройки (сокращение общего количества признаков) эффективно проводить классификацию признаков DoS- или DDoS-атаки [17]. В то же время, как показало исследование для построения алгоритмов многоуровневой классификации более эффективно использовать рекуррентные нейронные сети (Recurrent Neural Network, RNN), которые показали себя достаточно продуктивными в использовании и для бинарной классификации [18]. Также при решении указанной проблемы рассматривается адаптивный подход выбора признаков [19-21], который базируется на специальных функциях признаков DoS- или DDoS-атак и оценки подмножества параметров на основе согласованности (Consistency based Subset Evaluation, CSE). Данные функции могут быть использованы для эффективного выделения признаков кибератак на основе простой архитектуры нейросети, например, многослойный персептрон (Multi-layer perceptron, MLP), при бинарной либо многоуровневой классификации [20, 21].

Таким образом, технический анализ позволяет использовать большое количество подходов по выделению признаков DoS- и DDoS-атак нейросетевыми алгоритмами. Также следует также указать на необходимость решения в рамках указанной задачи таких типичных задач обучения нейросетей как дельта реконструкции (Low Reconstruction Error, LRE), взрывающийся градиент (Exploding Gradient, EG), затухающий градиент (Vanishing Gradient, VG) и переобучение, что рассматривается как ***нерешенная часть*** вопроса в рамках общего исследования.

Целью данной работы является разработка методики построения, настройки и оптимизация нейросетевых алгоритмов для выделения признаков DoS- и DDoS-атак с целью уменьшения нагрузки на вычислительный ресурс аппаратно-программной платформы, увеличения точности машинного анализа сетевого трафика, а также уменьшения времени как машинного анализа, так и обучения нейронной сети.

1. Методы подготовки обучающей выборки и предварительной обработки данных

Важным этапом подготовки нейронного сетевого алгоритма отслеживания признаков потенциальных киберугроз является подготовка обучающей выборки. В случае выделения признаков DoS- и DDoS-атак типичными наборами, на основе которых может быть сформирована обучающая выборка, являются наборы CICIDS [12] и NSL-KDD [13]. Преимуществами данных наборов являются включение актуальных и адекватных данных, а также подробная классификация признаков (43 класса в обучающей выборке NSL-KDD и 84 класса в обучающей выборке CICIDS, включая высокоуровневые признаки), которые разделяются на признаки, соответствующие входному траффику и метки класса кибератаки (наличие и степень угрозы).

В рамках данного исследования предлагается построить обучающую выборку на основе следующего набора требований:

- анонимность входных данных, на основе которых строится выборка, с учетом конфиденциальности, при сохранении полезной нагрузки, позволяющей выявить признаки киберугрозы;
- полный охват различных подходов при осуществлении DoS- и DDoS-атак (так называемых векторов возможных киберугроз), такие как атака методом перебора (Brute Force, BF), атака на основе браузера, атака с DNS-усилением, атака со сканирование портов, а также бэкдор-атака и ее разновидности;
- полный охват различных протоколов, применяемых для передачи данных в рамках инфраструктуры сетевого сервиса, таких как HTTP, HTTPS, DNS, SMTP, SSH и другие активно используемые протоколы;
- полный охват уязвимых элементов инфраструктуры сетевого сервиса, такие как наборы пакетов входного трафика, данные и функции маршрутизации, коммутатора, хоста, многоадресной IP-рассылки либо широковещательного домена (Broadcast Domain, BD);
- полный учет факторов взаимодействия между сетью (сетями) сетевого сервиса и сетью (сетями в случае DDoS-атаки) производится кибератака, в рамках чего рассматривается также и взаимодействие между локальными и глобальными сетями;

- полный набор данных о конфигурации аппаратно-программной платформы, на которую производится кибератака, включая модемы, брандмауэры, коммутаторы, маршрутизаторы, рабочие станции серверного комплекса и операционные системы (Operational System, OS);
- полный набор данных о входном трафике, включая неоднородность сетевого трафика в соответствии с реальной задачей;
- полный набор признаков внешней киберугрозы с маркировкой отдельных информационных блоков;
- метаданные, которые включают документацию о конфигурации распределенной сети, операционной системы злоумышленника, сценарии кибератак и наборе уязвимых данных.

При этом целью предварительной обработки данных является подготовка обучающей выборки в соответствии с особенностями поставленной задачи и принципами построения нейросетевых алгоритмов. Соответственно в рамках предварительной обработки данных выполняется обновление набора, оптимизация системы классификации через отказ от ненужных признаков, кодирование данных с помощью меток, а также их нормализация, что в рамках данного исследования формализуется следующим образом:

1. Кодирование набора данных обучающей выборки при помощи меток. Данный этап включает в себя преобразования категориальных признаков, которые преобразуются в числовые показатели (Hot Encoding, HE) при помощи хеш-функции (Hash Function, HF) на основе n-аргумента, что позволяет представить 2^n возможных уникальных значений.
2. Удаление признаков, которые являются нерелевантными в отношении задачи анализа данных. На данном этапе удаляются неинформативные или принципиально недопустимые значения с целью увеличения эффективности выполнения нейросетевых алгоритмов машинного анализа.
3. Нормализация признаков. На этапе кодирования данных каждому признаку (Feature Attribute, FA) в рамках хеширования присваивается значение в пределах $FA_i \in [FA_i^-; FA_i^+]$, которые являются уникальными для каждого признака. На этапе нормализации, в свою очередь, каждый признак набора FA_i для всех $i \in [1; I]$ проводится стандартная процедура мин-макс масштабирования (MinMax).

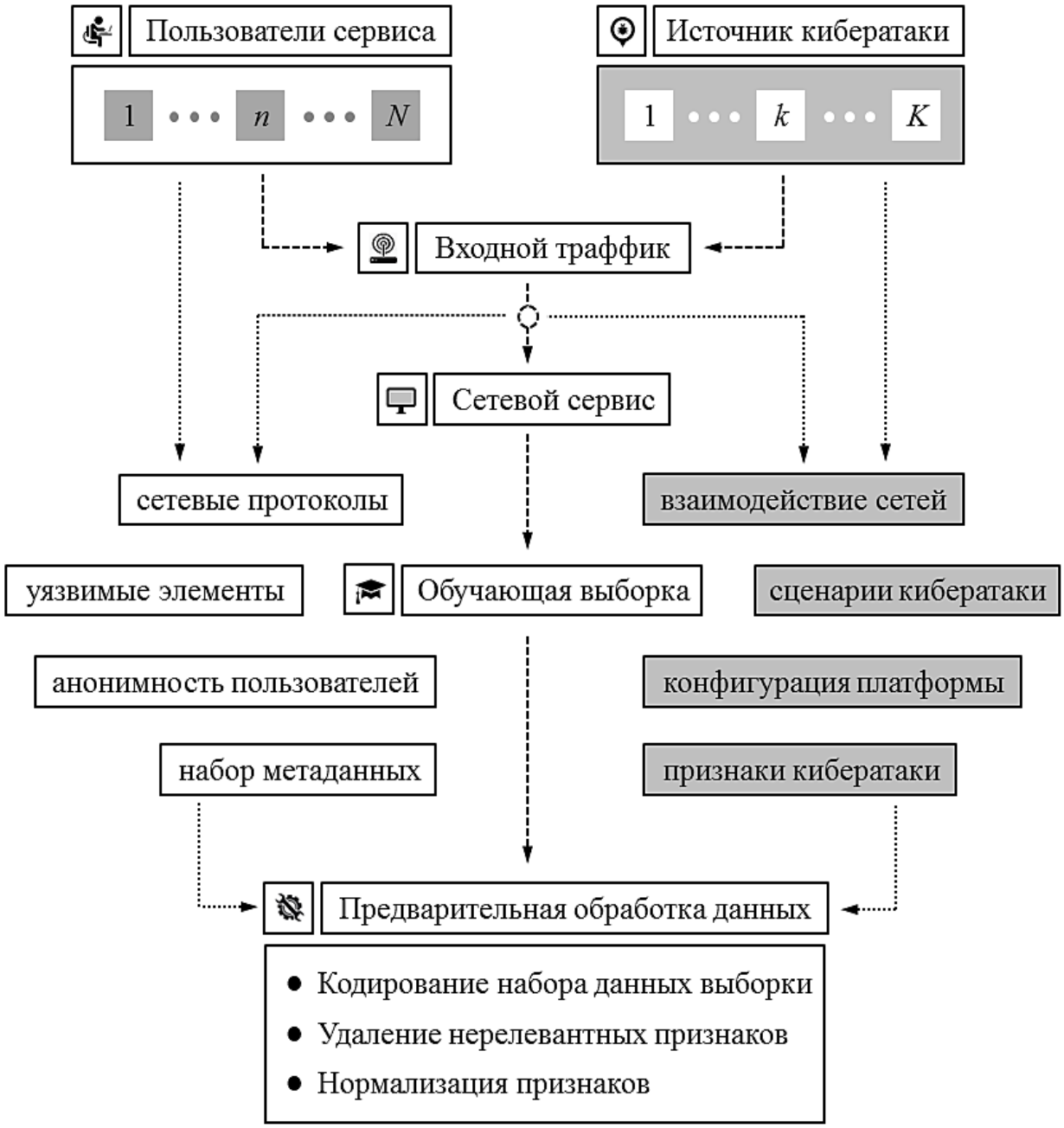

Рис. 1. Схема подготовки обучающей выборки и предварительной обработки данных

Математическое выражение для минимакс-масштабирования может быть представлено следующим образом:

$$\overline{FA_i} = \frac{FA_i - FA^i_-}{FA^i_+ - FA^i_-} \text{ для } \forall i \in [1; I]. \qquad (1)$$

Нормализация признаков позволяет оптимизировать использование обучающей выборки в нейронных сетевых алгоритмах машинной обработки данных входного трафика.

2. Организация нейросетевой архитектуры для выделения признаков DoS/DDoS-атак

Нейронные сетевые алгоритмы типа «автокодировщик» [17, 20, 21], которые рассматриваются в данном исследовании, базируются на полуавтоматическом машинном обучении (Semi-Supervised Machine Learning, S-SML), что позволяет упростить представление признаков и уменьшить нагрузку на вычислительный ресурс аппаратно-программного комплекса. Архитектура нейросети автокодировщик включает в себя один входной слой, один или несколько скрытых слоев, на уровне которых производится кодирование и декодирование данных, а также один выходной слой, на уровне которого проводится окончательный этап декодирования.

Представим выборку неразмеченных данных обучающего набора (Training Data, TD) как $\{TD_j\}$, а выходные данные, полученные через применение алгоритма обратного распространения (Back Propagation, BP) как $\{BP_j\}$, где $j \in [1; J]$, причем целью является соответствие $TD_j = BP_j$ для $\forall j$. Принцип работы нейронного сетевого

алгоритма на основе архитектуры типа автокодировщик, таким образом, включает в себя следующие этапы (рис. 2):

1. Входной набор данных обрабатывается функцией кодировщика (Encoder Function, EF) представленной через $F_E(TD_j)$, что позволяет уменьшить их размерность, получив соответствующий набор кодированных данных (Encoded Data, ED) — $\{ED_j\}$

2. Набор кодированных данных $\{ED_j\}$ проходит через этапы восстановления на основе функции декодировщика (Decoder Function, DF) — $F_D(ED_j)$, в результате чего на выходе нейронного сетевого алгоритма можно получить набор восстановленных данных (Reconstructed Data, RD) — $\{RD_j\}$ с размерностью соответствующей входному набору $\{TD_j\}$.

3. Процедура полуавтоматического обучения нейронного сетевого алгоритма типа автокодировщик при этом базируется на определении функции потерь $F_L(TD_j, RD_j)$, которая соотносит значения TD_j и RD_j для всех значений набора $j \in [1; J]$. Обучение базируется на уменьшении «потерь», что математически выражается $F_L(TD_j, RD_j) \to 0$.

Функция кодировщика нейронного сетевого алгоритма рассчитывается на основе весовых коэффициентов $\{w_j\}$, функции активации F_A и величины константного сдвига b:

$$ED_j = F_A(w_j \cdot TD_j + b) \text{ для } \forall j \in [1; J]. \quad (2)$$

В свою очередь, функция потерь рассчитывается через усреднение разницы между полным набором значений TD_j и RD_j:

$$F_L(TD_j, RD_j) = \frac{\sum_{j=1}^{J} |TD_j - RD_j|}{J}. \quad (3)$$

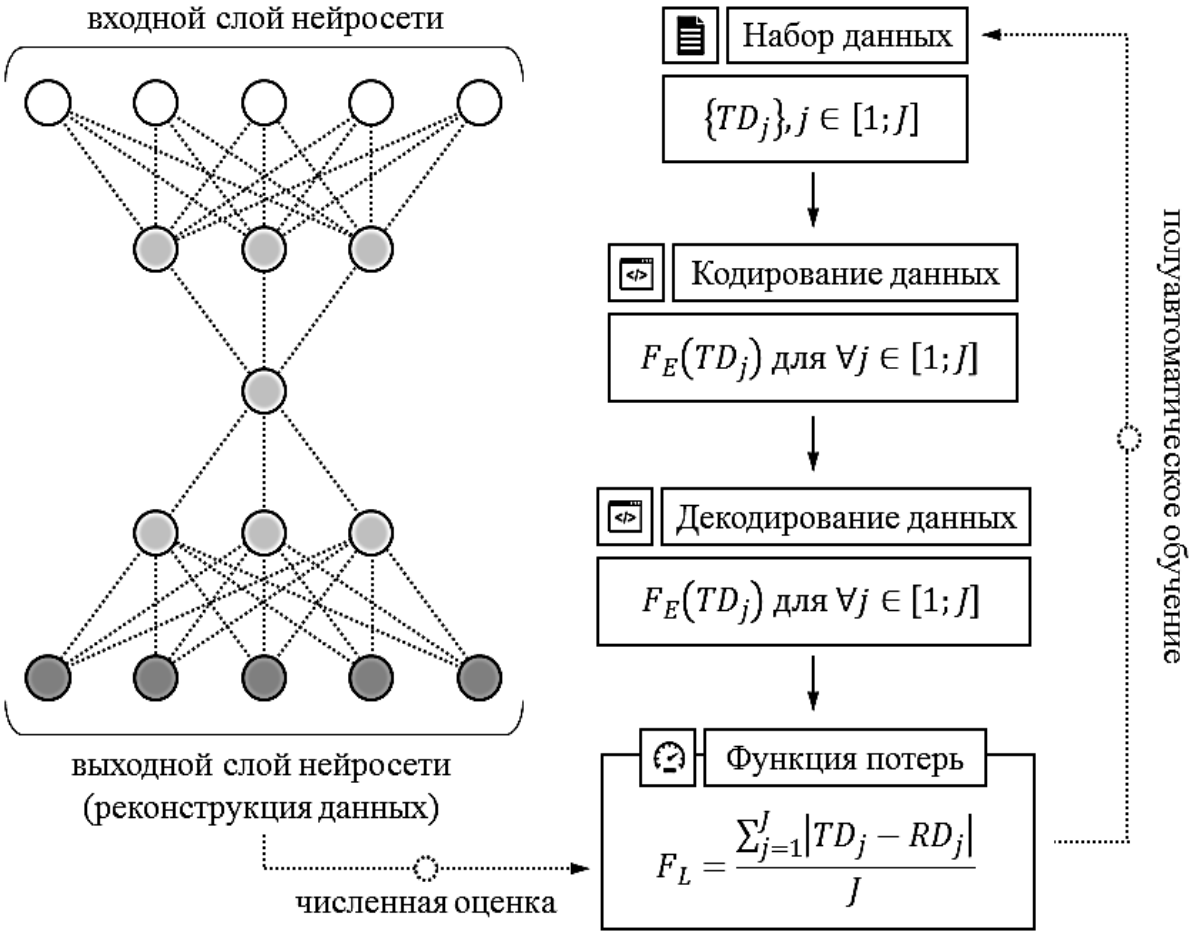

Рис. 2. Организация работы нейронного сетевого алгоритма модели автокодировщик при выделении признаков DoS-/DDoS-атак

Модель базового уровня, рассматриваемая в данном исследовании, базируется на двух обучающих выборках обнаружения сетевых вторжений, наборы которых разделяются в соответствии с типами кибератак. Предварительная обработка позволяет выделить признаки, имеющие категориальные значения, преобразуются в числовые значения с помощью прямого кодирования. Мин-макс масштабирование, в свою очередь, дает

возможность нормализовать данные числовые значения в диапазоне $[0;1]$. Полученная выборка используется для обучения автокодировщика, что по сравнению со стандартной обучающей выборкой (как заранее подготовленной, так и условно случайного набора) позволяет уменьшить время обучения и исключить проблему переобучения.

Далее полученный набор признаков, прошедший этап кодирования, подается на DNN для проведения классификации. Набор данных включает часть обучающего набора и часть тестового набора в соотношении 3:1. Схема DNN-автокодировщика отличается от стандартной схемы большим количеством скрытых слоев, «нейроны» которых полностью связаны с «нейронами» соседних слоев, что на уровне расширения функционала позволяет выделить высокоуровневые признаки киберугрозы. Это дает возможность однозначно отнести данную архитектуру к типу архитектуры нейронной сети с прямой связью, что указывает на снижение нагрузки на вычислительный ресурс аппаратно-программной платформы сетевого сервиса по сравнению с другими типами нейронных сетей глубинного обучения. Категории признаков, не подаваемые на вход автокодировщика с целью оптимизации процесса обучения, включают в себя идентификатор потока, IP-адрес источника внешней угрозы, порт аппаратной платформы и время проведения кибератаки. Это связано с тем, что данные категории являются информативными и их можно подать непосредственно на вход DNN.

Выводы

В результате проведенного анализа была разработана комплексная методология организации комплексной методологии построения нейронных сетевых алгоритмов выделения признаков DoS/DDoS-атаки, а также оптимизации данных алгоритмов в соответствии с целевыми показателями точности машинного анализа.

Таким образом, в рамках решения основной задачи было предложено использовать нейронную сетевую архитектуру типа «автокодировщик», что позволило разработать следующие подходы:

- схема подготовки обучающей выборки и предварительной обработки данных соответствующей выборки с целью оптимизации процедуры выделения и классификации признаков DoS/DDoS-атаки;
- схема организации основных принципов работы нейронного сетевого алгоритма модели «автокодировщик» при выделении и классификации признаков DoS/DDoS-атаки.

Данная методология благодаря построению математической модели, которая позволяет обобщить процессы машинного анализа может быть эффективно использована для решения широкого класса задач на уровне построения системы защиты сетевых сервисов от внешних угроз.

Список литературы / References

1. *Chiranjeevi H.S. & Manjula,K.S.* (2019). An text document retrieval system for University Support Service on a high performance distributed information system. 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA). [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/icccbda.2019.872576/ (дата обращения: 14.07.2022).
2. *Du Y., Liu J., Guan Z.,& Feng H.* (2018). A medical information service platform based on distributed cloud and Blockchain. 2018 IEEE International Conference on Smart Cloud (SmartCloud). [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/smartcloud.2018.00014/ (дата обращения: 14.07.2022).
3. *Li N. & Du Y.* (2013). Design and implementation of a cloud based Forensic Science Information System Model. 2013 International Conference on Cloud and Service Computing. [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/csc.2013.29/ (дата обращения: 14.07.2022).
4. *Qin Z., Zhou E., Ding Y., Zhao Y., Deng F. & Xiong H.* (2018). Data Service Outsourcing and privacy protection in Mobile internet. Data Service Outsourcing and Privacy Protection in Mobile Internet. [Электронный ресурс]. Режим доступа: https://doi.org/10.5772/intechopen.79903/ (дата обращения: 14.07.2022).
5. *Deepanshu Sharma M., Som S. & Khatri S.K.* (2017). Enhancing password security using cyclic group matrix. 2017 2nd International Conference on Telecommunication and Networks (TEL-NET). doi:10.1109/tel-net.2017.8343549.
6. *Hadaad N., Drury L. & Addie R.G.* (2015). Protecting services from security mis-configuration. 2015 International Telecommunication Networks and Applications Conference (ITNAC). doi:10.1109/atnac.2015.7366799.
7. *Varshney G. & Gupta H.* (2017). A security framework for IOT devices against wireless threats. 2017 2nd International Conference on Telecommunication and Networks (TEL-NET). doi:10.1109/tel-net.2017.8343548.
8. *Washburn T.* (2019). Cyber-attack. Thorndike Press, a part of Gale, a Cengage Company.
9. *Feng J., Chen Y., Summerville D., Ku W.S., Su Z.* Enhancing cloud storage security against roll-back attacks with a new fair multi-party non-repudiation protocol. In: Consumer Communications and Networking Conference (CCNC), 2011 IEEE. Pp. 521–522. IEEE (2011).
10. *Acharya S. & Pradhan N.* (2017). DDoS simulation and hybrid ddos defense mechanism. International Journal of Computer Applications, 163(9), 20–24. [Электронный ресурс]. Режим доступа: https://doi.org/10.5120/ijca2017913736/ (дата обращения: 14.07.2022).
11. *Gupta B.B. & Dahiya A. (*2021). Fundamentals of ddos attack: Evolution and challenges. Distributed Denial of Service (DDoS) Attacks, 1–18. [Электронный ресурс]. Режим доступа: https://doi.org/10.1201/ 9781003107354-1/
12. *Sharafaldin I., Habibi Lashkari A. & Ghorbani A.A.* (2019). A detailed analysis of the CICIDS2017 Data Set. Communications in Computer and Information Science, 172–188. [Электронный ресурс]. Режим доступа: https://doi.org/10.1007/978-3-030-25109-3_9/ (дата обращения: 14.07.2022).

13. *Thomas R. & Pavithran,D.* (2018). A survey of intrusion detection models based on NSL-KDD data set. 2018 Fifth HCT Information Technology Trends (ITT). [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/ctit.2018.8649498/ (дата обращения: 14.07.2022).
14. *Meena G. & Choudhary R.R.* (2017). A review paper on IDS classification using KDD 99 and NSL KDD dataset in Weka. 2017 International Conference on Computer, Communications and Electronics (Comptelix). [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/comptelix.2017. 8004032/ (дата обращения: 14.07.2022).
15. *Dong S., Abbas K. & Jain R.* (2019). A survey on distributed denial of service (ddos) attacks in SDN and Cloud Computing Environments. IEEE Access, 7, 80813–80828. [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/access.2019.2922196/ (дата обращения: 14.07.2022).
16. *Praseed A. & Thilagam P.S.* (2019). DDoS attacks at the Application Layer: Challenges and research perspectives for safeguarding web applications. IEEE Communications Surveys & Tutorials, 21(1), 661–685. [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/comst.2018.2870658/ (дата обращения: 14.07.2022).
17. *Yousefi-Azar M., Varadharajan V., Hamey L. & Tupakula U.* (2017). Autoencoder-based feature learning for cyber security applications. 2017 International Joint Conference on Neural Networks (IJCNN). [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/ijcnn.2017.7966342/ (дата обращения: 14.07.2022).
18. *Yin C., Zhu Y., Fei J. & He X.* (2017). A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access, 5, 21954–21961. [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/access.2017.2762418/ (дата обращения: 14.07.2022).
19. *Yusof A.R., Udzir N.I., Selamat A., Hamdan H. & Abdullah M.T.* (2017). Adaptive feature selection for denial of services (DOS) attack. 2017 IEEE Conference on Application, Information and Network Security (AINS). [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/ains.2017.8270429/ (дата обращения: 14.07.2022).
20. *Al-Qatf M., Lasheng Y., Al-Habib M. & Al-Sabahi K.* (2018). Deep learning approach combining sparse Autoencoder with SVM for network intrusion detection. IEEE Access. 6. 52843–52856. [Электронный ресурс]. Режим доступа: https://doi.org/10.1109/access.2018.2869577/ (дата обращения: 14.07.2022).
21. *Binbusayyis,A. & Vaiyapuri T.* (2021). Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM. Applied Intelligence. 51(10). 7094–7108. [Электронный ресурс]. Режим доступа: https://doi.org/10.1007/s10489-021-02205-9/ (дата обращения: 14.07.2022).