UDC 004.896
IRSTI 28.23.29

https://doi.org/10.55452/1998-6688-2024-21-3-48-57

[1]Omarov Bauyrzhan.S.,
PhD student, ORCID ID: 0000-0002-9312-4429,
e-mail: bauyrzhanomarov01@gmail.com
[2]Auelbekov O.A.,
Cand. Phys.-Math. Sc., senior researcher, ORCID ID: 0000-0002-2903-9086,
e-mail: omirlan.auelbek@gmail.com
[3]*Kulambayev B.O.,
Cand. Tech. Sc., Associate Professor, ORCID ID 0009-0002-9279-6239,
*e-mail: bakhytzhan.kulambaev@gmail.com
[1]Omarov B.S.,
PhD, acting associate professor, ORCID ID: 0000-0002-8341-7113,
e-mail: batyahan@gmail.com

[1]Al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan
[2]Institute Information and Computational Technologies, 050000, Almaty, Kazakhstan
[3]Turan University, 050013, Almaty, Kazakhstan

IOT NETWORK INTRUSION DETECTION USING MACHINE LEARNING ON UNSW-NB15 DATASET

Abstract

This research presents a comprehensive investigation into the application of machine learning techniques for addressing the pervasive security challenges within Internet of Things (IoT) networks. With the exponential growth of interconnected devices, ensuring the integrity and confidentiality of data transmissions has become increasingly critical. In this study, we deploy and evaluate seven distinct machine learning methods tailored to the IoT network intrusion detection problem. Leveraging the rich and diverse UNSW-NB15 dataset, encompassing real-world network traffic scenarios, our analysis encompasses a thorough examination of both traditional and state-of-the-art algorithms. Through rigorous experimentation and performance evaluation, we assess the efficacy of these methods in accurately detecting and classifying various forms of network intrusions. Our findings provide valuable insights into the strengths and limitations of different machine learning approaches for enhancing the security posture of IoT environments, thereby facilitating informed decision-making for network administrators and cybersecurity practitioners.

Key words: IoT network, intrusion detection, IoT attack, machine learning, artificial intelligence.

Introduction

The advent of the Internet of Things (IoT) has ushered in a new era of connectivity, interconnecting countless devices and systems to enhance efficiency, convenience, and productivity across various domains [1]. However, the proliferation of IoT devices and the sheer volume of data they generate have raised significant concerns regarding security and privacy [2]. As these devices become increasingly integrated into critical infrastructure, homes, and industries, they become attractive targets for malicious actors seeking unauthorized access, data breaches, or disruption of services [3].

IoT intrusion detection has emerged as a pivotal defense mechanism against these threats, aiming to detect and mitigate potential intrusions or malicious activities within IoT ecosystems [4]. Machine learning methods analyze data through computational algorithms, extracting patterns and features indicative of network intrusions, and subsequently employ decision-making processes to classify and discern between benign and malicious activities within IoT network traffic [5]. By leveraging

ML algorithms, it becomes possible to develop proactive, adaptive, and efficient security measures capable of addressing the dynamic nature of IoT networks [6].

This research paper explores the intricate landscape of IoT attack detection applying machine learning methods. It delves into the various challenges associated with securing IoT ecosystems and highlights the significance of robust intrusion detection systems in safeguarding the IoT data [7]. Moreover, this paper offers a comprehensive analysis of the existing ML techniques applied to IoT security, shedding light on their strengths and limitations [8].

The goal of this research is to contribute to the ongoing discourse on IoT security by offering insights into the efficacy of ML-based intrusion detection systems. Through rigorous experimentation and evaluation, we aim to assess the efficiency of different ML algorithms in identifying and mitigating threats within IoT environments. This research intends to provide practical recommendations for enhancing the security posture of IoT deployments, with implications for industries, governments, and individuals [9].

In the subsequent sections, we will delve deeper into the foundations of IoT intrusion detection, review pertinent literature, discuss the methodologies employed in this research, present experimental results, and offer conclusions and future directions. In this paper, we anticipate that readers will gain a comprehensive understanding of the intricate interplay between IoT, machine learning, and intrusion detection, and how these elements collectively contribute to the broader landscape of IoT security.

Main provisions

The proposed research reveals that the application of anomaly detection techniques introduces a notable challenge related to the categorization of data. The task entails segregating network traffic into two discrete categories: regular and anomalous, thus constituting a binary classification problem [10]. This delineation is pivotal for discerning between typical network behavior and potentially malicious activities within the IoT environment.

In tackling this binary classification challenge, fundamental mathematical methodologies will be utilized to detect notable deviations present within the network traffic graph. This approach aims to effectively distinguish between normal and abnormal patterns, thereby enhancing the robustness of intrusion detection systems deployed in IoT networks. These techniques will enable us to detect and isolate instances of severe fluctuations or deviations that signify potential anomalies in the IoT network traffic. Equation (1) demonstrates mathematical model of the IoT intrusion detection process using machine learning.

$$S = \int_{t_1}^{t_2} \left| x'(t) \right| dt \tag{1}$$

The aggregate manifestation of potential deviations within the temporal span from t1 to t2 is encapsulated herein. Within the framework of a discretized function, the formulation is articulated as follows:

$$S = \sum_{t=t_1}^{t_2-1} \left| x(t+1) - x(t) \right| \tag{2}$$

In the following segment, machine learning methodologies are applied to unveil discrepancies inherent within the IoT network, followed by an assessment utilizing various measurement parameters customized to the dataset at hand.

Materials and Methods

The research process can be structured into a comprehensive three-phase framework, as illustrated in Figure 1. The initial stage entails a meticulous system modeling process, aimed at

delineating three discrete categories: normal operations, malfunctions, and potential attacks. This modeling endeavor serves as a cornerstone for comprehending the varied system behaviors across these diverse scenarios.

Moving on to the second phase, an extensive range of execution scenarios is systematically conducted. The primary objective here is to generate datasets that provide a comprehensive and nuanced representation of the system's performance across the aforementioned conditions – namely, normal, defective, and under attack [11]. This step involves the deliberate manipulation of variables and conditions to capture the full spectrum of system behavior.

In the third and final stage of the research process, the amassed datasets become invaluable assets for the evaluation of numerous supervised machine learning algorithms [12]. The focus of this phase centers on assessing the efficacy of these algorithms, particularly in addressing the classification challenge inherent in distinguishing between the different system states identified earlier [13]. Through this rigorous evaluation, the research aims to contribute insights into the suitability and performance of various machine learning techniques in the context of system state classification.

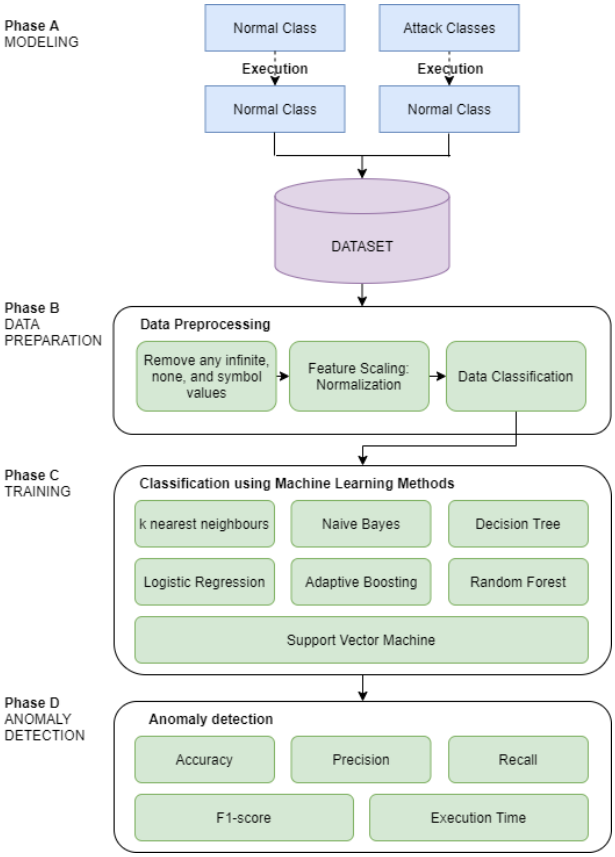

Figure 1 – Flowchart of the study

Dataset

The UNSW-NB15 dataset plays a pivotal role as a primary dataset within this research endeavor [14]. This dataset comprises a diverse and representative collection of network traffic data, encompassing a multitude of network-based attacks and normal network activities [15]. It is meticulously constructed to simulate real-world network traffic scenarios, making it an invaluable resource for assessing the performance of intrusion detection algorithms and machine learning models.

One of the main characteristics of the applied dataset is its categorization into various classes, including different types of attacks (e.g., denial-of-service attacks, intrusion attempts, and exploitation),

as well as benign or normal network traffic [16]. This categorization enables researchers to train and evaluate their models on a wide spectrum of network behaviors, facilitating the development of robust intrusion detection systems.

Moreover, the dataset provides a wealth of attributes and features, ranging from basic network flow statistics to more advanced protocol-specific characteristics. These attributes offer a comprehensive view of network traffic patterns, allowing researchers to employ a variety of ML methods to uncover hidden insights and patterns.

In this research, the UNSW-NB15 dataset serves as a foundational element, enabling the evaluation and validation of the machine learning approaches employed for IoT network anomaly detection. By utilizing this dataset, the research aims to harness its diversity and complexity to develop and assess effective intrusion detection models tailored to IoT environments, ultimately contributing to the enhancement of IoT security.

Evaluation Parameters

In this section, we provide an overview of the research topic and outline the key objectives and scope of the study. We introduce the importance of the subject matter, highlight existing gaps or challenges, and offer a roadmap for the subsequent sections of the paper [17–18].

Following the classification process, the outcomes can be categorized into four distinct types: TP means True Positive values, TN means True Negative values, FP means False Positive values, and FN means False Negative values. These outcomes are encapsulated within an error matrix [19].

In the evaluation of machine learning models, several essential metrics are employed to assess their performance, each offering valuable insights into their effectiveness and capabilities.

Accuracy is a fundamental evaluation metric used in machine learning to assess the performance of classification models [20]. It represents the proportion of correctly classified instances among the total number of instances in the dataset. A higher accuracy score indicates that the model has made fewer mistakes in its predictions, reflecting its ability to effectively distinguish between different classes. However, accuracy alone may not provide a complete picture of model performance, especially in imbalanced datasets where one class significantly outnumbers the others. In such cases, it is essential to consider additional metrics such as precision, recall, and F1 score to better evaluate the model's effectiveness.

$$accuracy = \frac{TP + TN}{P + N} \tag{3}$$

Precision focuses on the model's ability to correctly identify positive instances among those it classified as positive [21]. It is a crucial metric in scenarios where false positives are costly or undesirable, such as medical diagnoses.

$$preision = \frac{TP}{TP + FP} \tag{4}$$

Recall, also known as sensitivity or true positive rate, gauges the model's capability to identify all positive instances within the dataset [22]. High recall is essential when missing a positive instance could have significant consequences.

$$recall = \frac{TP}{TP + FN} \tag{5}$$

The F-score combines precision and recall into a single metric, striking a balance between the two [23]. It's particularly valuable when there's a trade-off between precision and recall, helping find an optimal threshold for classification.

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{6}$$

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are essential evaluation metrics used to assess the performance of binary classification models. The ROC curve illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate (1 – specificity) across various threshold values. A perfect classifier would have an ROC curve that reaches the top-left corner of the plot, indicating high sensitivity and low false positive rate across all thresholds.

The AUC-ROC metric quantifies the overall performance of the classifier by computing the area under the ROC curve. A higher AUC-ROC score indicates better discrimination capability of the model, with a value of 1 representing a perfect classifier and 0.5 representing random guessing. It provides a single scalar value to compare and rank different classifiers, making it a valuable tool for model selection and comparison.

The ROC-AUC metric is particularly useful in scenarios where class imbalance exists or when the costs of false positives and false negatives are not equal. It provides a comprehensive assessment of the model's ability to rank instances correctly, irrespective of the chosen threshold, making it a widely used metric in various fields, including healthcare, finance, and cybersecurity.

Results and Discussion

In the context of the burgeoning field of Internet of Things (IoT) security, the task of network intrusion detection is paramount for safeguarding interconnected devices and systems. The research presented in Table 1 delineates the efficacy of various machine learning models in identifying unauthorized access within an IoT network. This study meticulously evaluates the models based on four critical metrics: Accuracy, Precision, Recall, and F-score, which collectively offer a comprehensive view of each model's performance.

The K-Nearest Neighbours (KNN) algorithm demonstrates a commendable performance with an accuracy of 86.03%, indicating its proficiency in correctly identifying intrusion instances. Its precision and recall, closely aligned at 86.12% and 86.02% respectively, along with an F-score of 86.02%, suggest a balanced capability in both identifying true positives and minimizing false negatives.

The Naïve Bayes classifier, with an accuracy of 83.05%, showcases its potential despite being based on the assumption of feature independence. The slight elevation in precision (83.79%) over recall (83.09%) underscores its conservative nature in classifying an action as an intrusion, which is further substantiated by its F-score of 83.47%.

Decision Trees (DT) exhibit a modest performance with an 82.27% accuracy and similarly aligned precision and recall rates. This model's simplicity and interpretability do not compromise its effectiveness in the intrusion detection domain, as evidenced by an F-score of 82.31%.

Support Vector Machines (SVM) emerge as a robust contender with the highest recall of 87.80%, indicating superior sensitivity in detecting true positives. Its overall accuracy stands at 87.21%, complemented by a precision of 87.18% and an F-score mirroring its precision, which highlights its strength in managing unbalanced classes inherent to intrusion detection tasks.

Logistic Regression and AdaBoost algorithms show notable accuracies of 86.08% and 87.54%, respectively, with AdaBoost slightly edging out in precision, recall, and F-score metrics. These outcomes underscore the adaptability of ensemble methods like AdaBoost in enhancing prediction accuracy through the combination of multiple weak learners.

The Random Forest model, an ensemble of Decision Trees, registers the highest accuracy (87.82%) and F-score (87.65%) among the evaluated models. This denotes its exceptional capability in handling the complexity and variability of IoT intrusion datasets, benefiting from both the robustness of ensemble learning and the depth of decision trees.

In summary, the Random Forest algorithm stands out as the most effective model for IoT network intrusion detection, demonstrating superior performance across all evaluated metrics. This analysis underscores the critical role of machine learning in fortifying IoT networks against sophisticated intrusion attempts, advocating for the adoption of advanced models like Random Forest for enhanced security measures in IoT ecosystems.

Table 1 – Obtained results in IoT network intrusion detection problem

Machine Learning Model	Accuracy	Precision	Recall	F-score
KNN	86.03	86.12	86.02	86.02
Naïve Bayes	83.05	83.79	83.09	83.47
DT	82.27	82.34	82.19	82.31
SVM	87.21	87.18	87.80	87.18
Logistic Regression	86.08	86.54	86.85	86.48
AdaBoost	87.54	87.24	87.31	87.60
Random Forest	87.82	87.85	87.84	87.65

The graph in Figure 2 provides a comprehensive comparison of the performance of various machine learning models on the task of IoT network intrusion detection, using four key metrics: Accuracy, Precision, Recall, and F-score.

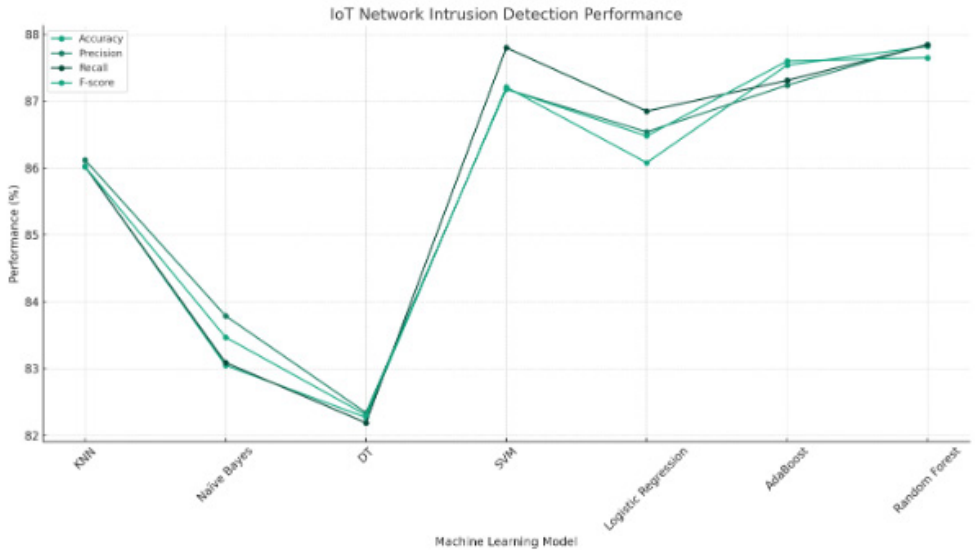

Figure 2 – Machine learning models in IoT network anomaly detection.

Random Forest stands out as the most effective model, demonstrating the highest scores across all metrics. This indicates its superior capability in correctly identifying both positive and negative instances of network intrusions, with minimal false positives and negatives. Its leading performance, particularly in Recall (87.84%) and Precision (87.85%), suggests it is highly reliable in identifying true intrusion cases without mistakenly flagging normal behavior as intrusive.

AdaBoost follows closely, with notably high scores, especially in F-score (87.60%), highlighting its effectiveness in balancing precision and recall. This suggests that AdaBoost is also highly capable of distinguishing between intrusion and non-intrusion instances, making it a strong candidate for IoT security applications.

SVM (Support Vector Machines) shows a distinct advantage in Recall (87.80%), the highest among all models, indicating its strength in identifying most true positive cases. However, its Precision and F-score, while competitive, do not outperform Random Forest or AdaBoost, suggesting that while it is excellent at detecting intrusions, it may have a slightly higher rate of false positives.

Logistic Regression and KNN (K-Nearest Neighbours) exhibit solid performance, with their metrics closely aligned. They both show a balanced trade-off between Precision and Recall, making them reliable choices for intrusion detection, albeit not as optimal as Random Forest or AdaBoost.

Naïve Bayes and DT (Decision Trees), while still effective, score lower compared to the other models. Naïve Bayes, despite its simplicity and fast computation, shows a limitation in precision and recall compared to more sophisticated models. Decision Trees present a foundational approach with decent performance but are outshined by their ensemble counterparts, such as Random Forest and AdaBoost, which leverage multiple trees for improved accuracy and generalization.

Overall, the graph illustrates the nuanced strengths and weaknesses of each model in the context of IoT network intrusion detection. Random Forest and AdaBoost emerge as the most promising models, offering a robust blend of high accuracy, precision, recall, and F-score, making them highly suitable for protecting IoT networks against intrusions.

Conclusion

In conclusion, this research endeavors to address the challenge of IoT network intrusion detection through the application of various machine learning methodologies. By employing a diverse set of classifiers including KNN, Naïve Bayes, Decision Tree, Support Vector Machine (SVM), Logistic Regression, AdaBoost, and Random Forest, we have conducted a comprehensive evaluation of their performance on the developed dataset. The results highlight the efficacy of Random Forest and AdaBoost in achieving high accuracy, precision, recall, and F-score values, signifying their suitability for intrusion detection tasks in IoT networks. Conversely, Naïve Bayes and Decision Tree models demonstrate comparatively lower performance across all evaluated metrics. These findings underscore the significance of selecting appropriate machine learning algorithms tailored to the intricacies of the dataset to enhance detection capabilities within IoT environments. Moreover, the study contributes valuable insights into the comparative analysis of machine learning techniques for intrusion detection, offering guidance for network administrators and cybersecurity practitioners in selecting optimal solutions to mitigate security threats. Future research directions may involve exploring ensemble methods or deep learning approaches to further enhance the accuracy and robustness of intrusion detection systems in IoT networks, addressing evolving cybersecurity challenges in the digital landscape.

REFERENCES

1 Ahmad M., Riaz Q., Zeeshan M., Tahir H., Haider S.A., & Khan M.S. Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using UNSW-NB15 data-set. EURASIP Journal on Wireless Communications and Networking, 2021, no. 1, pp. 1–23.

2 Aleesa A., Younis M. O. H. A. M. M. E. D., Mohammed A.A., & Sahar N. Deep-intrusion detection system with enhanced UNSW-NB15 dataset based on deep learning techniques. Journal of Engineering Science and Technology, 2021, vol. 16, no.1, pp. 711–727.

3 Fuat, T. Ü. R. K. Analysis of intrusion detection systems in UNSW-NB15 and NSL-KDD datasets with machine learning algorithms. Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 2023, vol. 12, no. 2, pp. 465–477.

4 Kabir M.H., Rajib M.S., Rahman, A. S. M. T., Rahman M.M., & Dey S.K. Network Intrusion Detection Using UNSW-NB15 Dataset: Stacking Machine Learning Based Approach. In 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), 2022, February, pp. 1–6, IEEE.

5 Sahar N., Mishra R., & Kalam S. Deep learning approach-based network intrusion detection system for fog-assisted iot. In Proceedings of international conference on big data, machine learning and their applications: ICBMA 2019, 2021, pp. 39–50, Springer Singapore.

6 Fathima A., Khan A., Uddin M.F., Waris M.M., Ahmad S. Sanin C., & Szczerbicki E. Performance Evaluation and Comparative Analysis of Machine Learning Models on the UNSW-NB15 Dataset: A Contemporary Approach to Cyber Threat Detection. Cybernetics and Systems, 2023, pp. 1–17.

7 Hossain Z., Sourov M.M.R., Khan M., & Rahman P. Network Intrusion Detection using Machine Learning Approaches. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2021, November, pp. 438–442. IEEE.

8 Sharma B., Sharma L., Lal C., & Roy S. (2023). Anomaly based network intrusion detection for IoT attacks using deep learning technique. Computers and Electrical Engineering, no.107, p. 108626.

9 Saheed Y.K., Abiodun A.I., Misra S., Holone M.K., & Colomo-Palacios, R. A machine learning-based intrusion detection for detecting internet of things network attacks. Alexandria Engineering Journal, 2022, vol. 61, no. 12, pp. 9395–9409.

10 Sarhan M., Layeghy S., & Portmann M. Towards a standard feature set for network intrusion detection system datasets. Mobile networks and applications, 2022, pp. 1–14.

11 Baich M., Hamim T., Sael N., & Chemlal Y. Machine Learning for IoT based networks intrusion detection: a comparative study. Procedia Computer Science, 2022, no. 215, pp. 742–751.

12 Shareena J., Ramdas A., & AP H. Intrusion detection system for iot botnet attacks using deep learning. SN Computer Science, 2021, vol. 2, no. 3, pp. 1–8.

13 Zhao R., Gui G., Xue Z., Yin J., Ohtsuki T., Adebisi B., & Gacanin H. (2021). A novel intrusion detection method based on lightweight neural network for internet of things. IEEE Internet of Things Journal, 2021, vol. 9, no. 12, pp. 9960–9972.

14 Baniasadi S., Rostami O., Martín D., & Kaveh M. A novel deep supervised learning-based approach for intrusion detection in IoT systems. Sensors, 2022, vol. 22, no. 12, p. 4459.

15 Yin Y., Jang-Jaccard J., Xu W., Singh A., Zhu J., Sabrina F., & Kwak J. IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset. Journal of Big Data, 2023, vol.10, no. 1, pp. 1–26.

16 Kumar S., & Pathak N.K. Evaluation Of Machine Learning Algorithms For Intrusion Detection Utilizing UNSW-NB15 Dataset. Journal of Pharmaceutical Negative Results, 2022, pp. 4819–4832.

17 Al-Ambusaidi M., Yinjun Z., Muhammad Y., & Yahya A. ML-IDS: an efficient ML-enabled intrusion detection system for securing IoT networks and applications. Soft Computing, 2024, vol. 28, no. 2, pp. 1765–1784.

18 Kumar N., & Sharma S. A Hybrid Modified Deep Learning Architecture for Intrusion Detection System with Optimal Feature Selection. Electronics, 2023, vol. 12, no. 19, p. 4050.

19 Ahmad Z., Shahid Khan A., Wai Shiang C., Abdullah J., & Ahmad F. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, 2021, vol. 32, no. 1, e4150.

20 Hammad M., Hewahi N., & Elmedany W. T - SNERF: A novel high accuracy machine learning approach for Intrusion Detection Systems. IET Information Security, 2021, vol. 15, no. 2, pp. 178–190.

21 Almomani O., Almaiah M.A., Alsaaidah A., Smadi S., Mohammad A.H., & Althunibat A. Machine learning classifiers for network intrusion detection system: comparative study. In 2021 International Conference on Information Technology (ICIT), 2021, July, pp. 440–445. IEEE.

22 Rashid M., Kamruzzaman J., Imam T., Wibowo S., & Gordon S. A tree-based stacking ensemble technique with feature selection for network intrusion detection. Applied Intelligence, 2022, vol. 52, no. 9, pp. 768–9781.

23 Abdulla A.R., & Jameel N.G.M. A Review on IoT Intrusion Detection Systems Using Supervised Machine Learning: Techniques, Datasets, 2023.

[1]Омаров Бауыржан.С.,
докторант, ORCID ID: 0000-0002-9312-4429,
e-mail: bauyrzhanomarov01@gmail.com
[2]Әуелбеков Ө.А.,
ф.-м.-ғ.к., а.ғ.қ., ORCID ID: 0000-0002-2903-9086,
e-mail: omirlan.auelbek@gmail.com
[3]*Куламбаев Б.О.,
т.ғ.к., қауымдастырылған профессор, ORCID ID: 0009-0002-9279-6239,
e-mail: bakhytzhan.kulambaev@gmail.com
[1]Омаров Б.С.,
PhD, доцент м.а., ORCID ID: 0000-0002-8341-7113,
e-mail: batyahan@gmail.com

[1]Әл-Фараби атындағы Қазақ ұлттық университеті,
050040, Алматы қ., Қазақтан
[2]Ақпараттық және Есептеу технологиялары институты,
050000, Алматы қ., Қазақстан
[3]«Тұран» университеті, 050013, Алматы қ., Қазақстан

UNSW-NB15 ДЕРЕКТЕР ЖИЫНТЫҒЫНДА МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП ИНТЕРНЕТ ЗАТТАР ЖЕЛІСІНЕ ЕНУДІ АНЫҚТАУ

Аңдатпа

Бұл зерттеу жұмысы интернет заттар (IoT) мәнмәтінінде желілік ауытқуларды анықтау мәселесін шешу үшін әртүрлі машиналық оқыту әдістерінің тиімділігін зерттейді. Бағалау параметрлерінің кең жиынтығын, соның ішінде дәлдік, прецизиондылық, қайтарымдылық, F1 бағасы, оқыту уақыты және қабылдаушының жұмыс сипаттамаларының (ROC) талдауын пайдалана отырып, алты түрлі машиналық оқыту әдісін жүйелі түрде салыстырылады. Алынған нәтижелер логистикалық регрессияның тәжірибелік қолданылу мүмкіндігін ерекше айқындайды, бұл оның теңдестірілген жұмыс сипаттамаларына байланысты сенімді таңдау екенін көрсетеді. Логистикалық регрессия желілік ауытқуларды анықтауда жоғары дәлдікті ғана емес, сонымен қатар оқыту уақытының айтарлықтай қысқарғанын да көрсетеді, бұл оны әсіресе ауытқуларға уақытылы жауап беретін шешуші нақты қолданбалар үшін қолайлы етеді. Бұл зерттеу Интернет заттары желілерінің қауіпсіздігі мен тұтастығын жақсарту, желілік ауытқуларды анықтаумен байланысты күрделі мәселелерді шешу және Интернет заттарының киберқауіпсіздігінің дамып келе жатқан ландшафтында осы әдістемелердің практикалық өзектілігін көрсету үшін машиналық оқыту әдістерін қолдану туралы құнды ақпарат береді.

Тірек сөздер: Интернет заттар, басқыншылық, шабуылды анықтау, машиналық оқыту, жіктеу.

[1]Омаров Бауыржан С.,
докторант, ORCID ID: 0000-0002-9312-4429,
e-mail: bauyrzhanomarov01@gmail.com
[2]Әуелбеков Ө.А.,
к.ф.-м.н., с.н.с., ORCID ID: 0000-0002-2903-9086,
e-mail: omirlan.auelbek@gmail.com
[3]*Куламбаев Б.О.,
к.т.н., ассоциированный профессор, ORCID ID: 0009-0002-9279-6239,
e-mail: bakhytzhan.kulambaev@gmail.com
[1]Омаров Б.С.,
PhD, и.о. доцента, ORCID ID: 0000-0002-8341-7113,
e-mail: batyahan@gmail.com

[1]Казахский национальный университет имени аль-Фараби,
050040, г. Алматы
[2]Институт информационных и вычислительных технологий,
05000, г. Алматы
[3]Университет «Туран», 050013, г. Алматы

ОБНАРУЖЕНИЕ ВТОРЖЕНИЙ В СЕТЬ ИНТЕРНЕТА ВЕЩЕЙ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ НА ОСНОВЕ НАБОРА ДАННЫХ UNSW-NB15

Аннотация

В данной исследовательской работе исследуется эффективность различных методов машинного обучения для решения задачи обнаружения сетевых аномалий в контексте сред Интернета вещей (IoT). Используя разнообразный набор параметров оценки, включая точность, прецизионность, отзыв, оценку F1, время обучения и анализ рабочих характеристик приемника (ROC), систематически сравниваются шесть различных методов машинного обучения. Полученные результаты подчеркивают практическую применимость логистической регрессии, которая является надежным выбором благодаря своим сбалансированным эксплуатационным характеристикам. Логистическая регрессия не только демонстрирует высокую точность обнаружения сетевых аномалий, но и значительно сокращает время обучения, что делает ее особенно подходящей для реальных приложений, где своевременное реагирование на аномалии имеет решающее значение. Это исследование дает ценную информацию о применении методов машинного обучения для повышения безопасности и целостности сетей Интернета вещей, решения сложных задач, связанных с обнаружением сетевых аномалий, и подчеркивает практическую значимость этих методологий в меняющемся ландшафте кибербезопасности Интернета вещей.

Ключевые слова: Интернет вещей, вторжение, обнаружение атак, машинное обучение, классификация.

Article submission date: 13.02.2024