KOKAND UNIVERSITY HERALD SCIENTIFIC E-JOURNAL

QO‘QON UNIVERSITETI XABARNOMASI
KOKAND UNIVERSITY HERALD
ВЕСТНИК КОКАНДСКОГО УНИВЕРСИТЕТА

ISSN: 2181-1695 OAK: 01-08/1819/6 www.herald.kokanduni.uz

IOT NETWORK INTRUSION DETECTION SYSTEM USING MACHINE LEARNING TECHNIQUES

Meliboev Azizjon Ikromjon o‘g‘li
Teacher, Kokand University

MAQOLA HAQIDA

Qabul qilindi: 24-iyun 2024-yil
Tasdiqlandi: 26-iyun 2024-yil
Jurnal soni: 11
Maqola raqami: 34
DOI: https://doi.org/10.54613/ku.v11i11.972

KALIT SO‘ZLAR/ Ключевые слова/ keywords

IoT, IDS, Machine learning, Data science, data analysis, review, platform

ANNOTATION

The proliferation of Internet of Things (IoT) devices has transformed various industries by providing smart and automated solutions. However, the extensive connectivity and diverse nature of IoT devices have also introduced significant security challenges, particularly in terms of network intrusion. This paper explores the development and implementation of an Intrusion Detection System (IDS) for IoT networks using Machine learning techniques. The proposed IDS aims to detect and mitigate various cyber threats by analyzing network traffic and identifying anomalous patterns indicative of intrusions. This research contributes to the field of IoT security by providing a robust and scalable intrusion detection solution that leverages the power of machine learning

Introduction: The rapid adoption of IoT technologies has led to an unprecedented increase in the number of connected devices, enhancing convenience and efficiency across different sectors. However, the heterogeneity and resource-constrained nature of IoT devices make them vulnerable to various security threats, including unauthorized access, data breaches, and network attacks. Traditional security measures are often inadequate in addressing these challenges, necessitating the development of advanced IDS tailored for IoT environments. This paper presents a comprehensive approach to designing an IoT Network Intrusion Detection System using machine learning techniques to identify and respond to security threats in real-time. In the deployment scenario once malicious traffic is identified, IDS notifies firewalls or intrusion prevention systems. According to methodologies, IDS for IoT environment is divided into two methods which are Signature-detection and Anomaly-detection. Signature detection works with pre-defined signatures and filters. This technique can inspect the defined intrusions effectively while an undefined attack record is not well determined. In contrast, the anomaly-detection technique relies on heuristic methods to detect the undefined attack behavior. That means when the system identifies a difference from benign traffic patterns then this traffic count as network intrusions. We used anomaly-detection way by using Machine learning techniques such as Support Vector Machines (SVM), Decision Trees, Random Forests.

Related works: Previous research has highlighted the limitations of conventional IDS in IoT environments due to the unique characteristics of IoT networks. Studies have shown that machine learning techniques, such as classification, clustering, and anomaly detection, offer promising solutions for enhancing IDS performance. Various machine learning algorithms, including Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks, have been employed to detect network intrusions. This section reviews existing methodologies and their effectiveness in addressing IoT-specific security challenges.

Azizjon et al. provides implementation the Deep learning models such as CNN, LSTM, RNN, GRU by using sequential data in a prearranged time range as a malicious traffic record for developing the IDS. The benign and attack records of network activities are classified, and a label is given for the supervised-learning method. They applied their models to three different benchmark data sets which are KDDCup ’99, NSL-KDD, UNSW NB15 to show the efficiency of Deep learning approaches.

Gyamfi et al. presents a comprehensive review of state-of-the-art network intrusion detection systems (NIDS) and security practices for IoT networks. They have analyzed the approaches based on MEC platforms and utilizing machine learning (ML) techniques.

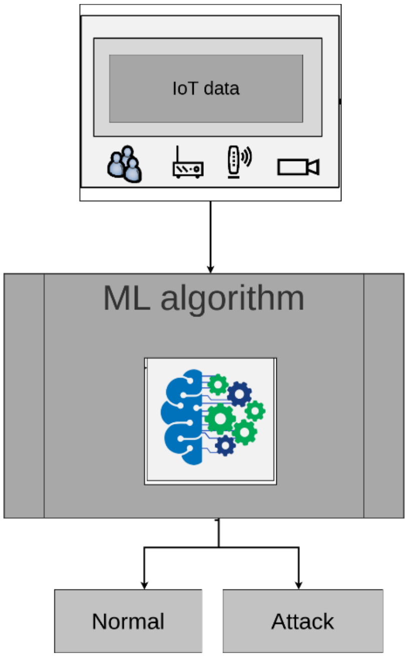
Chaabouni et al. focus on network intrusion detection systems (NIDSs). Their paper reviews existing NIDS implementation tools and datasets as well as free and open-source network sniffing software. Then, it surveys, analyzes, and compares state-of-the-art NIDS proposals in the IoT context in terms of architecture, detection methodologies, validation strategies, treated threats, and algorithm deployments. This paper will be useful for academia and industry research, first, to identify IoT threats and challenges, second, to implement their own NIDS and finally to propose new smart techniques in IoT context considering IoT limitations

The system's IoT has vulnerabilities in the controller, which is a crucial component and highly prone to various threats. Ismail et al. identified several failure points in the IoT environment, with one of the key vulnerabilities being the communication between the control and data planes. Attackers can exploit these vulnerabilities using "Botnets" to carry out saturation attacks, denial of service (DoS) or distributed denial of service (DDoS) attacks, and man-in-the-middle attacks, thereby exhausting the switch controller's bandwidth. DDoS attacks are classified into three main types: application-layer attacks, resource-draining attacks, and volumetric attacks. Application-layer attacks are sophisticated, targeting specific services with minimal bandwidth usage while gradually depleting network resources, making them difficult to detect. Examples include attacks on the Hypertext Transfer Protocol (HTTP) and Domain Name System (DNS). Resource-draining attacks exploit vulnerabilities in network layer protocols to make servers inaccessible, such as the TCP-SYN Flood which depletes the targeted machine's resources. Volumetric attacks aim to saturate network bandwidth by exploiting weaknesses in Layer 3 and Layer 4 protocols, executing attacks like ICMP, UDP, and TCP-SYN floods.

Methodology: The proposed IDS leverages machine learning algorithms to analyze network traffic and detect intrusions. The system architecture comprises three main components: feature extraction, data preprocessing and intrusion detection.

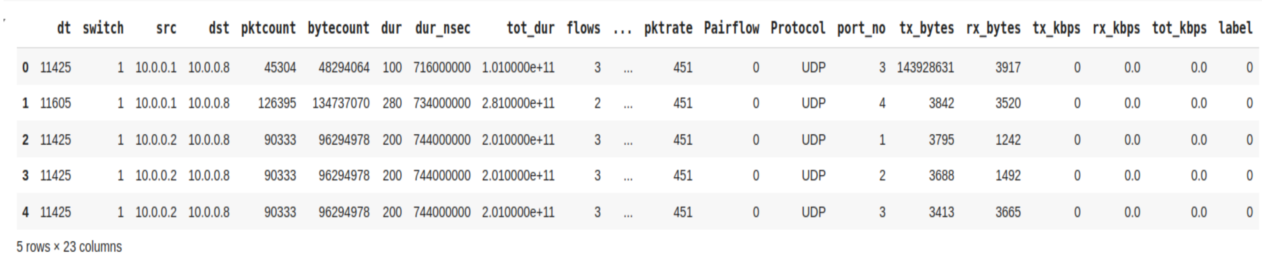
We present the stages of our proposed method for identifying DDoS attacks in IoT system. We developed a Machine learning solution to detect threats within the SDN environment, as illustrated in Figure 1. Our datasets comprise both Benign and DDoS attack data flows, with various features serving as input for our deep learning models. Detailed descriptions of the datasets are provided in the following subsection. During the training and evaluation phase, we preprocess the data by addressing missing values, performing one-hot encoding, scaling, and normalization. We then train several deep learning models, including SVM, Random Forest and Decision tree, using our dataset. These trained models are evaluated with testing data. In the final step, we can classify Benign and DDoS flows based on the trained models.

Figure 1. Model architecture.



Dataset description: Network traffic data is collected from IoT devices using network sniffers and stored in a centralized repository. The dataset includes various types of network packets, including benign and malicious traffic, to train and test the IDS. We used two recent DDoS dataset CICDDoS 2019. However, these dataset does not have features specifically tailored for IoT environments. The dataset, CICDDoS 2019, was chosen for its inclusion of network flow features and comprehensive labeling of data as either attack or benign. It comprises 80 network traffic features, extracted and computed for both benign and DDoS flows, and contains 97,831 benign instances and 333,540 attack instances. This dataset is publicly available and multiclass as shown in Figure 2.

Figure 2. Statistical information of data



	dt	switch	src	dst	pktcount	bytecount	dur	dur_nsec	tot_dur	flows	...	pktrate	Pairflow	Protocol	port_no	tx_bytes	rx_bytes	tx_kbps	rx_kbps	tot_kbps	label
0	11425	1	10.0.0.1	10.0.0.8	45304	48294064	100	716000000	1.010000e+11	3	...	451	0	UDP	3	143928631	3917	0	0.0	0.0	0
1	11605	1	10.0.0.1	10.0.0.8	126395	134737070	280	734000000	2.810000e+11	2	...	451	0	UDP	4	3842	3520	0	0.0	0.0	0
2	11425	1	10.0.0.2	10.0.0.8	90333	96294978	200	744000000	2.010000e+11	3	...	451	0	UDP	1	3795	1242	0	0.0	0.0	0
3	11425	1	10.0.0.2	10.0.0.8	90333	96294978	200	744000000	2.010000e+11	3	...	451	0	UDP	2	3688	1492	0	0.0	0.0	0
4	11425	1	10.0.0.2	10.0.0.8	90333	96294978	200	744000000	2.010000e+11	3	...	451	0	UDP	3	3413	3665	0	0.0	0.0	0

5 rows × 23 columns

Feature Extraction: Relevant features are extracted from the raw network traffic data to represent the characteristics of each network packet. Features include packet size, protocol type, source and destination IP addresses, and timing information. Feature selection techniques are employed to identify the most significant features contributing to intrusion detection.

Intrusion Detection

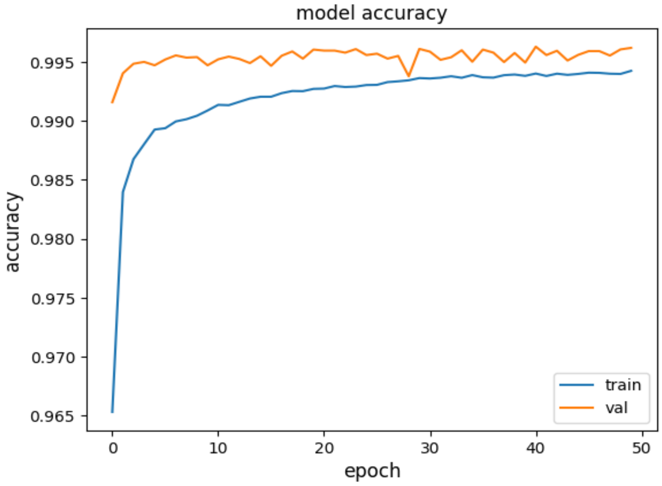
Machine learning algorithms are trained on the labeled dataset to distinguish between normal and malicious traffic. Various algorithms, including SVM, Random Forest, and Neural Networks, are evaluated for their accuracy, precision, recall, and F1-score. The trained models are then deployed to monitor real-time network traffic and detect potential intrusions.

Experimental Results and Discussion

The performance of the proposed IDS is evaluated using a publicly available IoT network dataset. The dataset is split into training and testing sets, and the models are assessed based on their detection accuracy and false positive rate. Comparative analysis of different machine learning algorithms is presented to identify the most effective approach for IoT intrusion detection.

The results demonstrate that machine learning techniques significantly enhance the detection accuracy of IDS in IoT networks. The choice of features and the quality of the training dataset play crucial roles in the system's performance. The paper also discusses the challenges associated with implementing machine learning-based IDS in resource-constrained IoT environments and suggests potential solutions. This plot is a confusion matrix for a SVM model, with an overall accuracy of 0.996 as shown in Figure. The matrix shows the performance of the model in classifying instances into two categories: "normal" and "attack."

Figure 3. Accuracy curve of SVM.



The SVM model has high accuracy in predicting attack traffic (0.99) but shows some misclassification of normal traffic as attack as illustrated in Figure 4. There are very few instances where attack traffic is misclassified as normal (0.012). The imbalance in the top-right cell suggests that normal traffic is often misclassified as attack, which could indicate a need for further tuning of the model to reduce false positives.

Figure 4. Confusion matrix of SVM

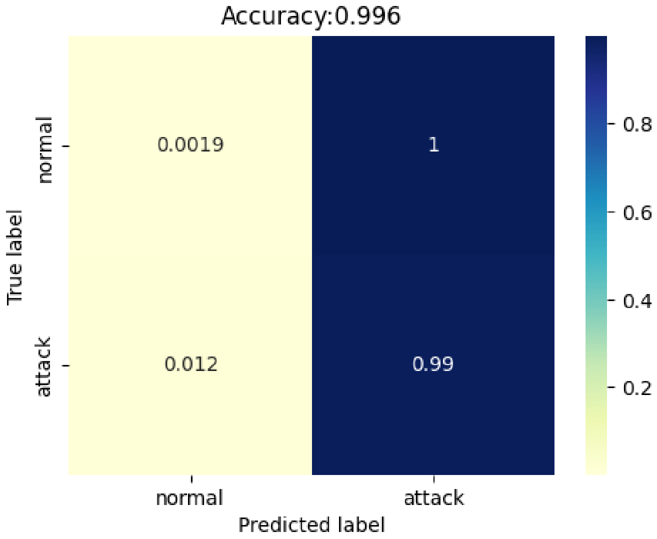


Figure 5 shows the training and validation accuracy of a Random Forest model over 50 epochs. Here's a detailed breakdown of what it illustrates:

Figure 5. Accuracy curve of Random forest

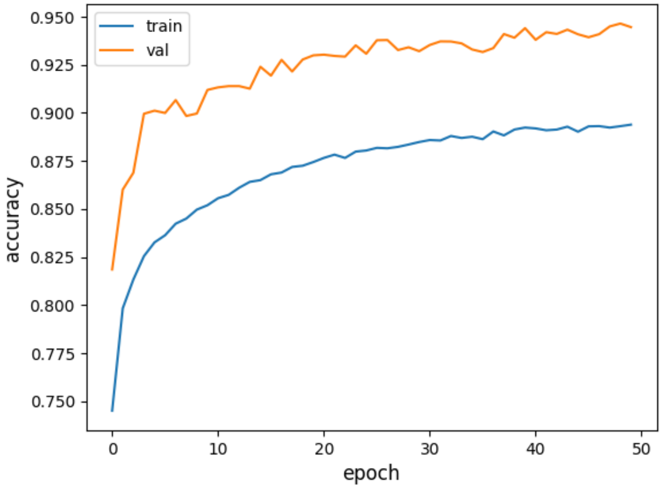
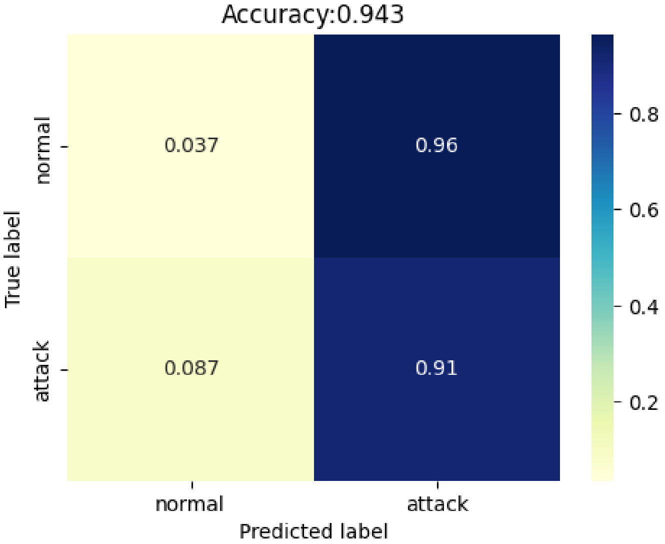


Figure 6 is a confusion matrix plot, typically used to evaluate the performance of a classification model. The model has high accuracy (0.943), but the confusion matrix indicates a significant number of normal cases being misclassified as attacks (FP = 0.96). This could suggest a bias in the model towards predicting attacks or an imbalance in the dataset. The model also has a reasonably good rate of detecting attacks correctly (TP = 0.91) and a lower rate of missing actual attacks (FN = 0.087).

Figure 6. Confusion matrix of Random forest



Conclusion: This paper presents an effective approach to developing an IoT Network Intrusion Detection System using machine learning techniques. The proposed system successfully identifies and mitigates various security threats in IoT networks, contributing to enhanced security and reliability of IoT deployments. Machine learning model SVM shows outperforms rather that other models. Future work will focus on optimizing the system for real-time detection and exploring advanced machine learning techniques, such as deep learning, for further improvements.

References:

1. Meliboev, A., Alikhanov, J., & Kim, W. (2022). Performance evaluation of deep learning based network intrusion detection system across multiple balanced and imbalanced datasets. *Electronics*, *11*(4), 515.

2. Gyamfi E, Jurcut A. Intrusion Detection in Internet of Things Systems: A Review on Design Approaches Leveraging Multi-Access Edge Computing, Machine Learning, and Datasets. Sensors (Basel). 2022 May 14;22(10):3744. doi: 10.3390/s22103744. PMID: 35632153; PMCID: PMC9143513.

3. N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac and P. Faruki, "Network Intrusion Detection for IoT Security Based on Learning Techniques," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2671-2701, thirdquarter 2019, doi: 10.1109/COMST.2019.2896380.

4. Azizjon, M., Jumabek, A., & Kim, W. (2020, February). 1D CNN based network intrusion detection with normalization on imbalanced data. In *2020 international conference on artificial intelligence in information and communication (ICAIIC)* (pp. 218-224). IEEE.

5. Hayotjon o‘g‘li, T. X., & Ikromjon o‘g‘li, M. A. (2023). BIG DATE TIZIMI HAQIDA UMUMIY TASNIF VA TUSHUNCHA. *QO ‘QON UNIVERSITETI XABARNOMASI*, 1281-1284.

6. Sarhan, Mohanad, Siamak Layeghy, Nour Moustafa, and Marius Portmann. NetFlow Datasets for Machine Learning-Based Network Intrusion Detection Systems. In *Big Data Technologies and Applications: 10th EAI International Conference, BDTA 2020, and 13th EAI International Conference on Wireless Internet, WiCON 2020, Virtual Event, December 11, 2020, Proceedings* (p. 117). Springer Nature.

7. Al-Rimy, B. A. S., Maarof, M. A., & Shaid, S. Z. M. (2018). Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions. Computers & Security, 74, 144-166.

8. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176.

9. Doshi, R., Apthorpe, N., & Feamster, N. (2018). Machine learning DDoS detection for consumer internet of things devices. In 2018 IEEE Security and Privacy Workshops (SPW) (pp. 29-35). IEEE.

10. Hodo, E., Bellekens, X., Hamilton, A., Dubouilh, P. L., Iorkyase, E., Tachtatzis, C., & Atkinson, R. (2016). Threat analysis of IoT networks using artificial neural network intrusion detection system. In 2016 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6). IEEE.

11. Kandasamy, P., Krishnan, G., & Chandrasekaran, K. (2020). A review on machine learning and deep learning techniques for intrusion detection system in IoT environment. International Journal of Computer Sciences and Engineering, 8(3), 90-97