2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) | 978-1-6654-7436-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAC3N56670.2022.10074106

Network Intrusion Detection System using ML

Raghav Kumar, School of Multidisciplinary Studies , REVA University, Bangalore

kumarraghav233@gmail.com

Abdul Haq Nalband, Assistant professor, School of Multidisciplinary Studies, REVA University, Bangalore

abdulhaq@reva.edu.in

***Abstract*—Technologies are making our life easier and simple, but it has both positive and negative effect. Many new methods of cybercrimes are coming which cannot be solved using earlier conventional method like using firewalls, antivirus, old ML algorithms. In recent years every device whether it's hardware or software is being connected with IOT. Hence, there is huge growth in data also their privacy is huge concern for industries. In this model, we are implementing Network Intrusion detection system using Machine learning algorithms which would resolve security problems using KNN, SVM, LR, RF, DT and Gaussian NB with greater efficiency. Our system uses both supervised and unsupervised machine learning techniques. Both misuse and Anomaly based detection to detect malware and viruses, our system is capable to detect both known and unknown attacks. In case of misuse detection system known attacks are being easily identified using a database where list of all known attacks is available. If any attack happens on network system, then NIDS checks whether the attack is listed in dataset or not. If attack is known, then system administrator gets notified. If attack is unknown, then NIDS uses outlier detection to identify attack using several machine learning algorithms like clustering and other techniques. So, with the help of above-mentioned techniques attack is being detected. Our model improves the attack detection mechanism with high accuracy and less prediction time. It is better than previous conventional machine learning algorithms. Our model is broadly accepted in companies and organization. it is fulfilling the cyber security issue also threat prediction time of our model is quite improved and the prediction time is reduced as compared to previous model.**

***Keywords*—KNN, SVM, LR, RF, DT, NB**

I. INTRODUCTION

The heightening of cyber crimes resulting several problems and their impact could be seen on both social and economic growth of people and organization. Now a days every company is doing their work in both offline and online mode. Also, most of enterprises are working in online mode to provide better service and product. some important applications that are used by humans like banking application, banking websites where person share their Credit information like account number, CVV and other details. These are the reasons behind enhancement of cyber-attacks. Hence, these types of data are used by hackers to do cyber bulling and other criminal activities. There is lot of noise about Deep learning (TSDO) which is better than conventional TSO [1] in terms of IDS. As per Cisco survey 2017, more than 40 percent of enterprises had come up with DDoS type of cyber-attack. Recently we heard that Spotify faced DDoS attack also GitHub had suffered from it. So, recently many companies [2] faced cyber-attack and it's damaging our social, economic, and politically also. Normal citizen and enterprises are suffering from cyber attacks also old firewall and antivirus techniques are not enough to stop it, but NIDS technique can detect attacks in whole network traffic, it may be host based or network-based intrusion detection system and would notify the users. In recent years advance type of cyber attacks are happening which made conventional technique ineffective. So, machine learning based misuse and anomaly detection [3] come with high precision and better efficiency to detect cyber threats by analyzing network traffic pattern.

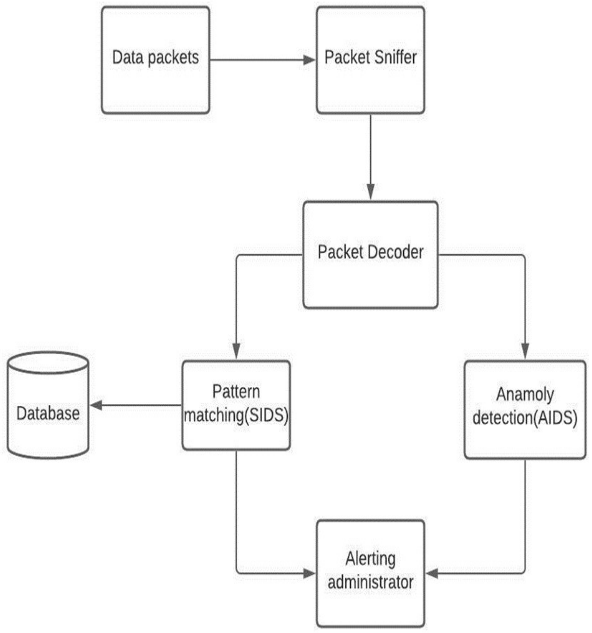

Fig 1. Network intrusion detection system diagram

NIDS (Network based intrusion detection system) monitors network traffic in every fraction of second. It monitors all the data packets which are moving in entire network system from one point to another and if they take the image of data packets in ideal state. If any mismatch found, then they compare the image of data packet with ideal state and notify the administrator. After detecting the attack, if type of attack is known then from set of datasets, it detects the attack. If new type of attack has happened on network infrastructure, then using anomaly detection technique it tries to find out attack pattern, in this process it uses several ML algorithms [4] like KNN, SVM, DT, LR, RF and Adaboost techniques. Most often hackers try new attack technique to penetrate network system. Our model also anatomize network traffic [5] in IOT devices and identify intruders in data packets. so, anomaly detection is most powerful technique for detecting anomalies in network.

II. Literature Survey

In the paper [1] "IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization" here, Deep learning and transient search optimization has been implemented to detect intrusion in IOT devices. In this model, we are presenting effective AI based intrusion detection system (IDS) in IOT systems. Deep learning and metaheuristics (MH) algorithms solved tedious engineering problems effectively. A deep learning-based CNN algorithm is used for feature extraction method. Also, in this model we are implementing a new variant of transient search optimization technique (TSO) for developing new feature selection method. Here, new variant of transient search optimization search algorithm (TSO) that is called TSODE. For it, our model is using Differential evolution technique (DE). TSODE is far better than conventional TSO techniques and, DE helped to elaborate the TSO technique.

Another paper named [2] "A novel attack detection scheme for the industrial internet of things using a lightweight random neural network" IOT products are used by independent customer for their personal use, IIOT are expanding in large scale. So, there should be high security and privacy of the system to give better service to the customers. In this model, a machine learning based lightweight random neural network (RANN) technique has been used to detect different types of attacks such as denial of service (Dos), probing, spying etc. Our RANN based ML technique model has several performance parameters such as precision, recall, f1score and accuracy. Also, our obtained output was collating with old artificial neural network (ANN), support vector machine (SVM) and other machine learning algorithms. RaNN technique find out attack with accuracy of greater than 99% and 34.51 milliseconds as prediction time.

Similarly, the paper [3] "A machine learning security framework for IOT systems" In this paper a machine learning (ML) based detection technique has been proposed for security aspects and all these process are related to IOT domain. The dominance of algorithms like software defined networking (SDN) and network function virtualization (NFV) are minimizing cyber threats as much as possible. Artificial intelligent framework analyzes the network pattern with the help of anomaly-based detection technique in IOT system. The proposed framework uses data mining, neural network, and supervised machine learning technique to obtain desired output. Evaluation of whole system done on three parameters like backpropagation technique, classification of entire model and rule association based on JRip Algorithm.

On further research a paper [4] called "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things" s. In this model an intrusion detection system is used to reduce the cyber-attacks. For particularly Botnet attacks, that happened on http and DNS protocols that is being used in internet of things. All these networking protocol's flow features are created based upon their characteristics. AdaBoost is a combined machine learning algorithm made by three different ML techniques like decision tree (DT), naïve bias (NB) and artificial neural network to identify malicious activity in IOT devices. So, with the help of different machine learning techniques and botnet detection tools, our model can identify the threats and would try to protect IOT devices as much as possible.

Finally, a paper named [5] "Intrusion Detection in Green Internet of Things: A Deep Deterministic Policy Gradient-Based Algorithm." This paper is all about intrusion detection system i.e. to identify attacks on system and to appropriate action so that losses would be minimized. Now a days, cyber intruder tries new ways to penetrate and take control of IOT devices. Also, it effects privacy system very badly. Intrusion detection system can detect attack in network traffic before any damage and would notify the administrator. Intrusion detection system is used to defense our network from outsider and insider attacks, also keeps our network safe and secure. Here, in this model an intrusion detection system based on deep reinforcement learning has been proposed which monitors the network in every fraction of second. IDS do some statistical features of prior network flow by monitoring network traffic and do traffic prediction. DDPG based IDS technique has been used to improve security of IOT devices by analyzing traffic flow and to detect intrusion. Also, traffic predictor is used to employ intrusion detection system. Our model has also checked the algorithm in detecting DDoS attack.

III. Network Intrusion detection System

NIDS is a type of intrusion detection system in which whole network traffic [4] is being monitored in every fraction of second to protect our IOT devices. In our network intrusion detection system, our model is using both signature and anomaly-based detection techniques. In the case of misuse detection system, there is a dataset where list of all known to attacks are available. If any known attack happened on the system, then it compares the attack with known attack dataset. Also, it determines the attack with certain pattern matching algorithms. Misuse detection has certain images of data packets in ideal state where there was not any malware or trojan in the system. NIDS takes the image of data packet in every fraction of second and compares the image of current data packets with ideal data packet image. If they found any mismatch in image, then they send the message to system administrator [5] to take appropriate action.

- If unknown attack is happened on the system, then our model cannot detect the threat using signature-based detection. So, in that case anomaly-based detection technique to detect threat. Now a days daily new type of attack is happening on the system, so anomaly detection is important.
- In anomaly detection system, machine learning algorithms is used to identify unknown attack, here

several machine learning algorithms are used to detect malware.

- k nearest neighbour, decision tree, logistic regression, random forest, naïve bias, and adaboost, also our model is using clustering techniques to identify threats.
- In the case of anomaly detection technique, we are using both supervised and unsupervised machine learning. In the case of supervised learning, our model need to train the model with labelled data set. So, that it can predict new attack type.
- .Our model also uses unsupervised learning technique to detect the attack type, also using clustering techniques to identify different type of attacks.

IV. METHODOLOGY

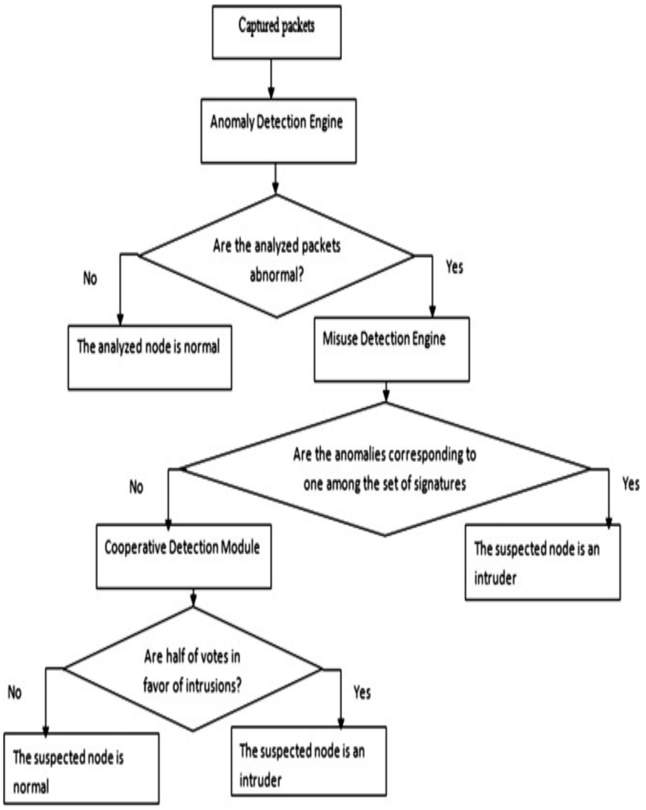
Captured packets
Anomaly Detection Engine
Are the analyzed packets abnormal?
No
Yes
The analyzed node is normal
Misuse Detection Engine
Are the anomalies corresponding to one among the set of signatures
No
Yes
Cooperative Detection Module
The suspected node is an intruder
Are half of votes in favor of intrusions?
No
Yes
The suspected node is normal
The suspected node is an intruder

Fig 2. Flowchart of network intrusion detection system

- In this paper, NIDS is analyzing the KDD datasets with respect to four classes which are basic, content, traffic, and host. So, in this way all databases are categorized.
- Our model has done analysis with the help of two evaluation techniques, and these are Detection rate (DR) and False Alarm Rate (FAR) for intrusion detection system (IDS).
- Based on our research work on KDD datasets, the help of all these four attributes on detection rate and false alarm rate are shown and these are used to increase the stability of data sets to get high detection rate and low false alarm rate.

A. Working Procedure

First, data packets are sent through the packet sniffer where all packets are monitored. Packet sniffer monitors the data packets flowing across the network system.

It can examine the network over larger internet also there are two ways to monitor network flow i.e., filtered, and unfiltered ways.

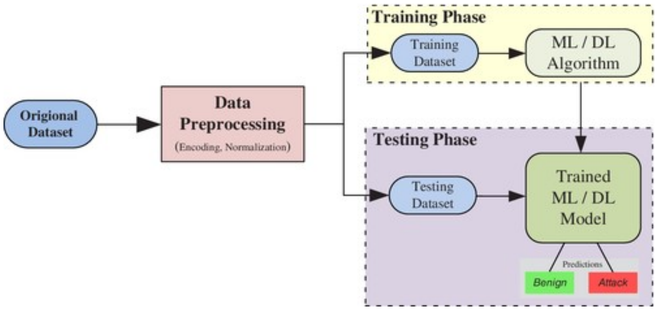

Fig 3. Training and Testing Phase of NIDS Model

In case of wireless network sniffer can scan one channel at a time but proposed model can enhance its capacity using multiple wireless interfaces.

Sometimes Companies also use sniffer to monitor their employees network system. It is used to clean network traffic and reduce malware infection.

After passing data packets through packet sniffer, now it is being passed through network intrusion detection system where if the type of attack is known then it falls under signature detection or else anomaly detection (for unknown attack).

If the attack pattern is known, then it tries to find pattern using database where the list of all known attacks is available.

If the intruder would have tried any new attack pattern, then anomaly detection comes into the picture where attacks are found out by supervised and unsupervised learning techniques.

Here, different ML algorithms like SVM, LR, RF, KNN, DT and AdaBoost techniques.

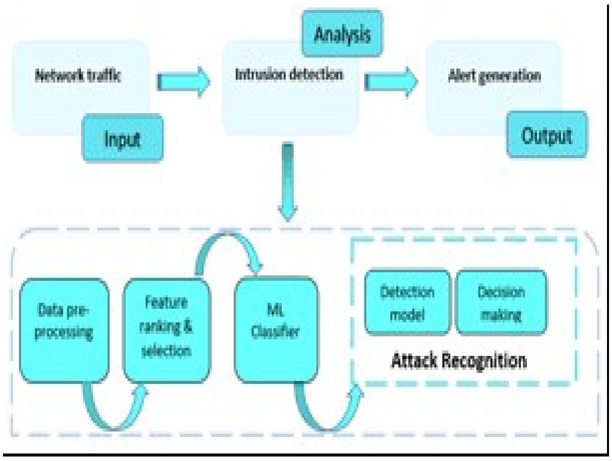

Fig 4. Detecting attack and alerting administrator in NIDS

All the outputs of model have been represented in the form of f1score, precision, recall, response time and efficiency and proposed model have compared our result with respect to various parameters and our result has shown tremendous result in terms of all mentioned parameters.

B. *Pros and Cons of NIDS:*

I. Pros of NIDS:

- It can detect malware across network infrastructure.
- NIDS monitors data packets in every fraction of second.
- NIDS uses both supervised and unsupervised ML techniques.
- Proposed model uses both misuse and anomaly-based detection technique.

II. Cons of NIDS:

- Intrusion prevention system (IPS) is missing in this model.
- Proposed model can only identify intruder and notify the admin but can't prevent the system.

C. ALGORITHM:

Step 1. Capturing the data packets.
Step 2. Packets are monitored by sniffer.
Step 3. Data packets are being decoded.
Step 4. Attacks are being determined.
Step 5. Known attacks are found by Misuse.
Step 6. Misuse uses set of Known datasets.
Step 7. Unknown attacks are found by anomaly.
Step 8. Anomaly uses ML Algorithms.
Step 9. Notify the system administrator.

V. RESULTS AND DISCUSSION

- In the area of cyber security, NIDS plays a significant role in protecting networks from attacks.
- In this work a Machine Learning based detection system has been proposed which is effectively better than previous conventional method.
- Our model is widely acceptable in industries, and it is for both insider and outsider attack.
- Our model is using python programming language to develop intrusion detection software and its accuracy and performance are high.
- To run our software, we are using python code and executing on JUPYTER Notebook which is open-source web application.

Our model has obtained our results in terms of following domains:

- Different types of attacks and their detection rate.
- Different networking protocols and their detection rates in terms of time and value.
- Different types of machine learning algorithms and their detection rates in terms of precision, recall, f1score, efficiency and performance.

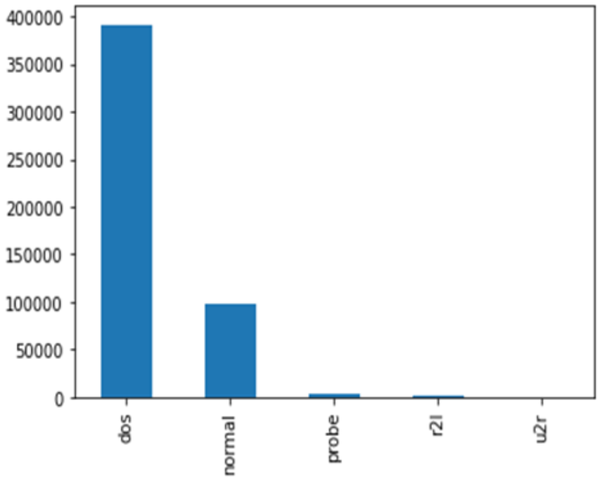

Fig 5. Different types of attack and no of their occurrence

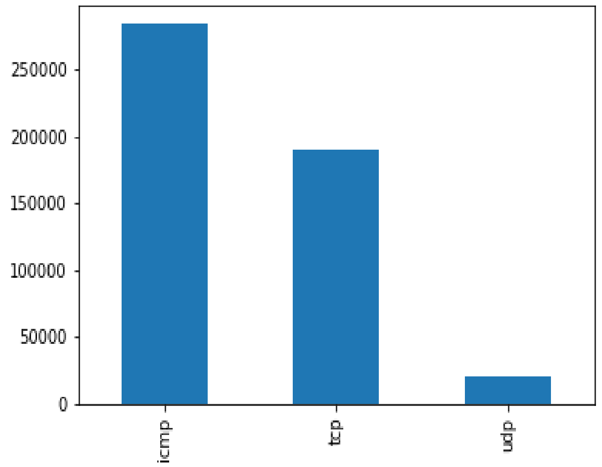

Fig 6. Protocol type: We notice that ICMP is the most present in the used data, then TCP and almost 20000 packets of UDP type.

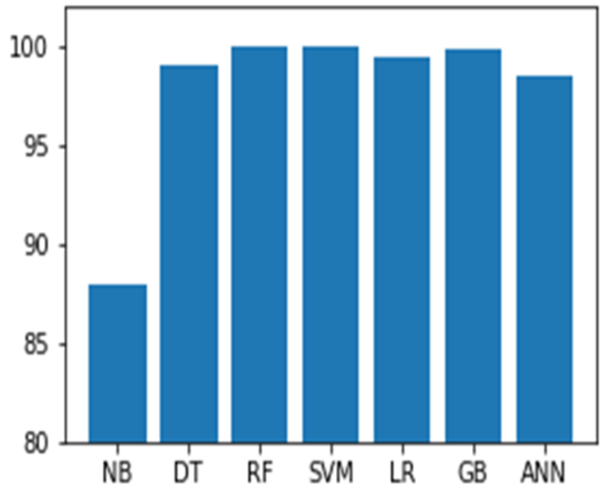

Fig 7. Training Accuracy of our NIDS Model

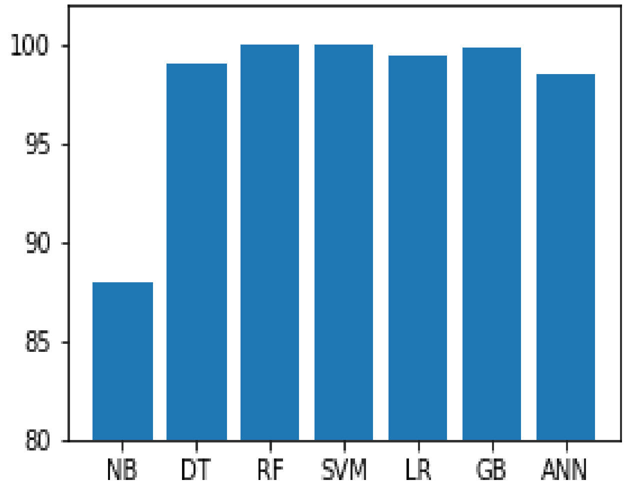

Fig 8. Testing Accuracy of Our NIDS Model

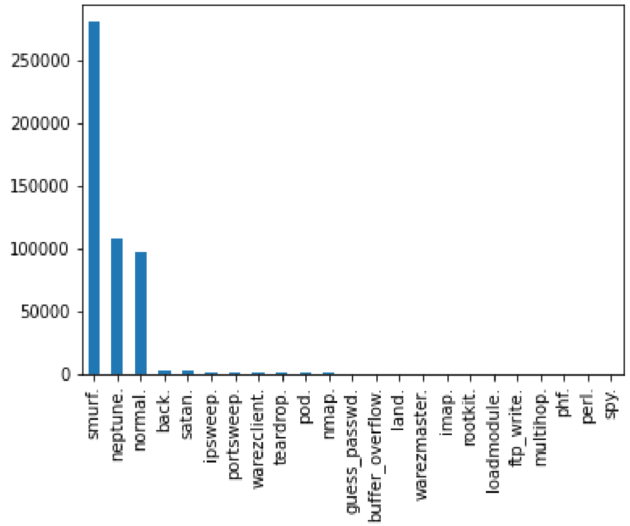

Fig 9. Attack Type (The attack types grouped by attack, it's what we will predict)

All the outputs of our model are represented in terms of diagram such as detecting different types of attack based on attack pattern. Attack detection in terms of different networking protocols.

Output of different ML algorithms are collected, and their results are shown as per the rate of detection.

VI. CONCLUSION

Our research on research on network intrusion detection system with the help of different machine learning algorithms like SVM, KNN, LR, RF, DT and gaussian naïve bayes are very impactful and it can detect intrusion with very less response time and better efficiency. In our software we are using both misuse and anomaly-based detection technique to detect intrusion.

Our main aim is to reduce the false message that generates during anomaly. To reduce losses as much as possible and try to protect system vulnerability, integrity, confidentiality, and availability are our main purpose. Our model is better than conventional techniques like firewalls, antiviruses software and other techniques also, our model is meeting to the user requirement.

VII. FUTURE SCOPE

It's quite amazing and challenging to explore our work from current industrial and social challenges to obtain safe and secure future of upcoming and current generation. In today's life where cyber threats are growing with the rapid growth of IOT devices.

Cyber-crime has been increasing tremendously and due to this there is huge concern of data that are generating from IOT devices. Many new intrusion detection mechanisms are coming to predict attacks.

Whenever any intruder tries to invade network then intrusion detection system notifies the administrator. Our model uses signature detection in the case of known attacks and for unknown attacks, anomaly detection system has been used.

Now a days hacking tools are become easy due to huge of amount of data that has been generating in different sectors. So, their privacy is huge concern. If our model is further improving with some more algorithms and techniques, then it would be more sophisticated and reliable.

ACKNOWLEDGEMENTS

I would like to thank Professor Abdul Haq who has been helping me throughout the whole research period. In addition, I also would like to thank Assistant Director Dr. Mohammad Riyaz Ahmad for his guidance and help throughout my work.

REFERENCES

[1] Fatani, A., Abd Elaziz, M., Dahou, A., Al-Qaness, M. A., & Lu, S. (2021). IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization. IEEE Access, 9, 123448-123464.

[2] Latif, S., Zou, Z., Idrees, Z., & Ahmad, J. (2020). A novel attack detection scheme for the industrial internet of things using a lightweight random neural network . IEEE Access, 8, 89337-89350.

[3] Bagaa, M., Taleb, T., Bernabe, J. B., & Skarmeta, A. (2020). A machine learning security framework for iot systems. IEEE Access, 8, 114066-114077.

[4] Moustafa, N., Turnbull, B., & Choo, K. K. R. (2018). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. IEEE Internet of Things Journal, 6(3), 4815-4830.

[5] Nie, L., Sun, W., Wang, S., Ning, Z., Rodrigues, J. J., Wu, Y., & Li, S. (2021). Intrusion Detection in Green Internet of Things: A Deep Deterministic Policy Gradient Based Algorithm. IEEE Transactions on Green Communications and Networking, 5(2), 778-788.

[6] M. Ahmed, A. N. Mahmood, and J. Hu, “A survey of network anomaly detection techniques”, Journal of Network and Computer Applications, vol. 60, 2016.

[7] Syarif I, Prugel Bennett A, Wills G., “Unsupervised clustering approach for network anomaly detection”, Networked Digital Technologies Communications in Computer and Information Science, vol. 293. Berlin Heidelberg: Springer, 2012, pp.135–45.

[8] A. Lakhina, M. Crovella and C. Diot, “Mining Anomalies Using Traffic Feature Distributions”, Proc. of ACM SIGCOMM, 2005.

[9] S. Novakov, C.-H. Lung, I. Lambadaris, Ioannis N. Seddigh, “Studies in applying PCA and wavelet algorithms for network traffic anomaly detection”, Proc. of IEEE 14th International Conference on High Performance Switching and Routing, 2013, pp. 185-190.

[10]L. Hu, T. Li, N. Xie, J. Hu, “False Positive Elimination in Intrusion Detection Based on Clustering”, Proc. of the 12th International Conference on Fuzzy Systems and Knowledge Discovery, 2012..

[11] C-F Tsai, Y-F Hsu, C-Y Lin, W-Y Lin, “Intrusion detection by machine learning: A review”, Journal on Expert Systems with Applications, vol. 36, 2009.

[12] M. Eslamnezhad and A-Y Varjani, “Intrusion Detection Based on MinMax K-means Clustering”, Proc. of the 7th International Symposium on Telecommunications, 2014.