Attack Detection in IoT using Machine Learning

Maryam Anwer
Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
maryam.anwer2@gmail.com

Shariq Mahmood Khan
Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
shariq@neduet.edu.pk

Muhammad Umer Farooq
Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
umer@neduet.edu.pk

Waseemullah
Department of Computer Science & IT
NED University of Engineer and Technology
Karachi, Pakistan
waseemu@neduet.edu.pk

Abstract-Many researchers have examined the risks imposed by the Internet of Things (IoT) devices on big companies and smart towns. Due to the high adoption of IoT, their character, inherent mobility, and standardization limitations, smart mechanisms, capable of automatically detecting suspicious movement on IoT devices connected to the local networks are needed. With the increase of IoT devices connected through internet, the capacity of web traffic increased. Due to this change, attack detection through common methods and old data processing techniques is now obsolete. Detection of attacks in IoT and detecting malicious traffic in the early stages is a very challenging problem due to the increase in the size of network traffic. In this paper, a framework is recommended for the detection of malicious network traffic. The framework uses three popular classification-based malicious network traffic detection methods, namely Support Vector Machine (SVM), Gradient Boosted Decision Trees (GBDT), and Random Forest (RF), with RF supervised machine learning algorithm achieving far better accuracy (85.34%). The dataset NSL KDD was used in the recommended framework and the performances in terms of training, predicting time, specificity, and accuracy were compared.

Keywords-cyber security; artificial intelligence; IoT; machine learning

I. INTRODUCTION

The Internet of Things (IoT) is probably the greatest modern advancement, considering its effect on our daily life, while the zones of its utilization are quickly expanding. In 2018, the quantity of IoT devices was roughly 28 billion. This amount is expected to touch 49.1 billion by 2022 and the showcase size of IoT is estimated to reach around $10 trillion by 2022. IoT is recognized as a method regarding suitable mechanisms that interconnect by servers, sensors, and various software. A city structure, is shown in Figure 1 which comprises of three main layers: fog, cloud, and terminal layer.

The data obtained from the IoT are saved on the Cloud Computing (CC) ecosystem which has progressively high-level processors and sufficient memory. The cloud layer has grown fast by the modern developments in IoT. Fog-to-things is created with a feasible clarification of those difficulties. In the fog layer, devices can experience some larger values of data basically given to the cloud layer, which decreases power damage, bandwidth, network traffic, and eliminates the data storage and communication challenges. In addition, it tries to accelerate the estimated method near the endpoint, facilitating some fast reply to the IoT-based urban use. There are two advantages of attack detection in the fog-to-things layer. Either the internet service provider or the network administrator can practice certain measures which can stop extensive destruction if these network attacks are recognized in the fog layer. Besides, this strategy does not prevent the regular daily experience for the people.

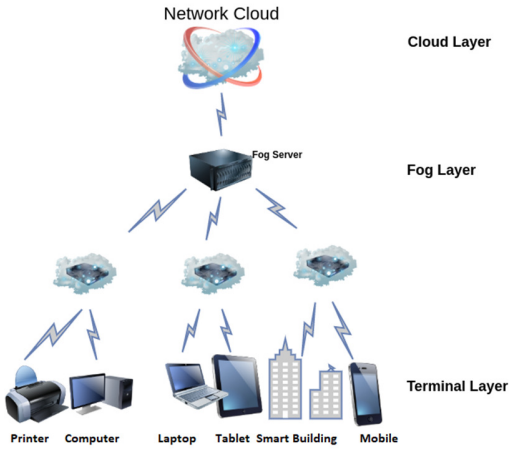

Fig. 1. Framework of a smart IoT-based city.

The model traces the web traffic which passes by every fog-to-things node. As fog-to-things connections resemble IoT devices, it will be more efficient to recognize these network attacks at the fog-to-things connections rather than at the cloud

Corresponding author: Shariq Mahmood Khan

layer. Immediate attack detection can inform the network controllers of the IoT devices of those attacks, which will then support them to evaluate and improve their systems. Artificial intelligence technology like Machine Learning (ML) will do the whole evaluation and send video pictures to people who can react speedily to solve troubles and maintain residents' safety. There are two types of attack detection: primarily signature-based or primarily anomaly-based. In the former, a primarily based solution fits the in-coming traffic closer to acknowledge attack/crime kinds in database whilst the latter checks the behavior of everyday traffic.

II. RELATED WORK

Many research studies in the application of ML have presently been presented in the domain, like object identification/recognition, pattern recognition, text processing, and image processing. In addition, much security research had been done using the Deep Learning (DL) approach [1]. Authors in [2] describe the expansion of big data and the evolution of IoT in a smart city. The author in [3] explains the evolution of CC and how big data have been engaged in the advancement of smart cities. He proposed a framework for managing big data for smart city purposes. The framework concentrates on difficulties related to smart cities for real-time decision planning. Many aspects and components of a smart city for upgrading the standard of the people are described in [4]. Authors in [5] suggest a platform design to secure a smart city facing cyber attackers. The structure is giving a warning DL model to identify attackers based on the user's data performance. In [6], resource administration methods of fog-computing are analyzed, well-systematic research in taxonomy is presented, and various features of resource administration, i.e. mass balancing, resource/device scheduling and allocation, job/task allocation, device/resource provisioning, and task offloading, are highlighted. The given resource management procedures are analyzed by estimating factors such as: Qos metrics, different researches, and applied methods. The benefits and hindrances of these approaches are compared.

Authors in [7] used the idea of an unknown and secure total plan (ASAS) in mist-based open distributed computing. In ASAS, the cloud gives advanced information about open cloud servers. When the ASAS is used, the fog gives devices to exchange information with PCS. Authors in [8] reported the advancements of remote sensor organization (WSN), correspondence innovation, and IoT innovation. Authors in [9] used ML techniques such as KNN, SVM, DT, Naïve Bayes, neural networks, and RF which can be applied in IDS. The authors compared ML models for multi and binary class combinations on the data set of Bot-IoT. Depending on these models they calculated the F1 score, recall, precision, and accuracy. The detection of attacks in FOG design was examined in [10], in which ML is compared with deep-learning neural networks working on an internet-available dataset.

Authors in [11] examined TCP SYN network attacks and authors in [12] introduced deep neural networks for attack detection in IoT systems. The self-adaptive identification method of the security index of the network was studied, performed risk assessment was conducted, and the system was mapped. Authors in [13] developed network NIDS based on the conception of DL. For attack detection, they implemented network intrusion detection system on fog node. Authors in [14] used a novel method that combines isolation forest and One Class Support Vector Machine (OCSVM) with an active learning method to detect attacks with no prior information. Authors in [15] used a two-stage approach combining a fast preprocessing or filtering method with a variation auto encoder using reconstruction probability. Authors in [16] performed a Distributed Denial of Service (DDoS) attack using the ping of death technique and detected it using RF algorithm by using the WEKA tool with classification accuracy of 99.76%. Authors in [17] proposed the detection of network dictionary attacks using a data set collected as flows based on a clustered graph. The results of the mentioned methods on the CAIDA 2007 data set give high accuracy for the model.

III. GAP ANALYSIS

These are some prefaced problems taken from earlier research.

- Worst performance of the detection of attacks on the fog layer.
- Feature selection decreasing the accuracy.
- Low accuracy of DoS, R2L, and U2R attack types.
- Execution of multiple classifier algorithms on reduced data sets
- False positive rate and false negatives rate is still in doubt.

IV. A FRAMEWORK TO SOLVE ATTACK DETECTION IN IOT USING MACHINE LEARNING

The proposed model for this research work is an ordinary huge organization or a smart city going through an increasing variety of IoT-associated cyber threats, such as heavy-obligation DDoS attacks, achieved with an enormous botnet, e.g. Mirai, which exploit default or weak passwords. The current research specializes in advanced attacks which can be primarily based on violations of organizational protection guidelines. Once completed, an attacker is permitted to take advantage of individuals who connect unauthorized styles of IoT devices to the smart town. The previous approaches have been used broadly because of their excessive detection accuracy and low fake alarms. However, they lack the capability of seizing novel attacks. On the other hand, anomaly detection detects new attacks, although it lacks accuracy. In both procedures, classical ML analysis has been used prominently. Popular devices gaining knowledge of algorithms are incapable to detect complex data breaches [18]. In this research, we examined different algorithms for the different sub-processes of the framework shown in Figure 2 [19].

A. Approaches to Solve Attack Detection using ML

There are six main approaches in ML:

- Supervised learning: In this, the data should be labeled like feeding a model with multiple examples of files and decide whether they are malware or not. Based on this data labeling [20], the model could decide on extra data. It is also called the task driven approach.

- Ensemble learning: It is an addition of label data like supervised learning while combining multiple models to solve the task.
- Unsupervised learning: In this learning, unlabeled data are used and the model marks them by itself based on the data properties. It is considered to be the more powerful and it usually finds anomalies in the data set [21]. This is also called the data-driven approach.
- Semi-super user learning: It tries to combine both supervised and unsupervised approaches when there is a data set with some labeled data [22].
- Reinforcement learning: This behavior should be used in a changing environment. It is also called the environment driven approach [23].
- Active learning: It works like a teacher who can help in correcting error and behavior in environmental changes [24]. It is a subclass of reinforcement learning.

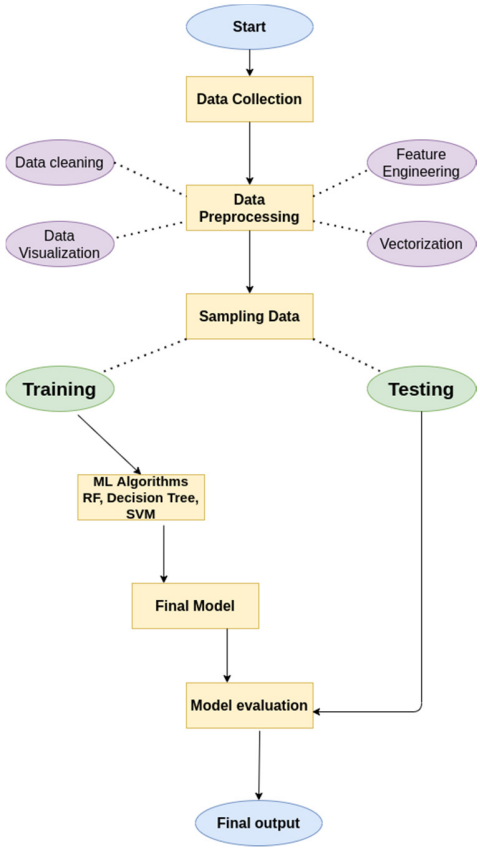

Fig. 2. A framework to solve attack detection in IoT using ML.

B. Attack Detection using ML Methods

In this part, the attack detection problems are studied by statistical classification of measurements using the implementation of ML. The spam filter in cyber security separates spam from different communications services. Spam is apparently the leading ML method applied in information security. The supervised learning labeled data method is usually used for classification. In our research, we used the Gradient Boosted DT, SVM, and RF classifications and the results were compared.

1) Support Vector Machine

It is the most popular and widely recognized technique. It can be used for regression, but mostly it is used in classification algorithms. In SVM, we sketch data items by the point in an n-dimensional area where n represents the considered features [25]. It creates a hyper plane and separates the data into classes [26].

2) Gradient Boosted Decision Tree

GBDT is an ensemble of DTs. GBDT is an ML algorithm which constructs vulnerable DTs through the boosting technique. For building the tree ensemble, we need to train over the algorithms on different samples. Unfortunately, we cannot train them on a single set. GBDT uses the present-day ensemble to predict the label of every instance, after which the results are compared with the accurate labeled data. It works on large datasets and has high predicting power [27].

3) Random Forest

RF [28] is based on random subspace, bagging, and uses CART DTs as base algorithm. It works on both regression and classification. The education is achieved in parallel. It injects randomness within the learning (testing and training), a process in which each tree isn't the same with the others. In predictions, each tree is combined, which reduces the variance of prediction and hence improves performance [29].

V. Experiments and Results

In this section, the dataset, which is applied for the experiment and for testing results, is described along with the performance metrics used for result comparison and the recommended model is reviewed by applying various selections and classifications. Three ML algorithms were applied for the evaluation of the given proposed model [30].

A. Dataset

The NSL KDD dataset was used in this research. This dataset is available in CSV and JSON files. We can use this for the model and the evaluation phase. The dataset is modifiable, extensible, and reproducible [31].

B. Proposed Method

Our research is a novel combination of several independent ML algorithms. In our framework, the first step is the dataset collection and analysis. In this process, the data were collected and observed deeply to analyze the types of data. In the data preprocessing step, the data were cleaned, visualized, and feature engineering was applied along with implemented vectorizations. Hence, the data were converted into feature vectors [32]. After the analysis of the NSL-KDD dataset, the attacks can be categorized into four principal classes:

- Unauthorized to remote (R2L)
- Denial-of-Service (DoS)
- Unauthorized to root super user privileges (U2R attack)

- Port scanning attack (Probe)

The details of each attack are shown in Figure 3.

TABLE I. TRAIN SET OF NSL-KDD

Type	Original records	Distinct records	Reduction rate
Attacks	3,925,640	262,178	93.3
Normal	972,782	812,814	16.44
Total	4,798,431	1,074,992	78.05

TABLE II. TEST SET OF NSL-KDD

Type	Original records	Distinct records	Reduction rate
Attacks	250,436	29,378	88.26
Normal	60,591	47,911	20.92
Total	311,027	77,289	75.15

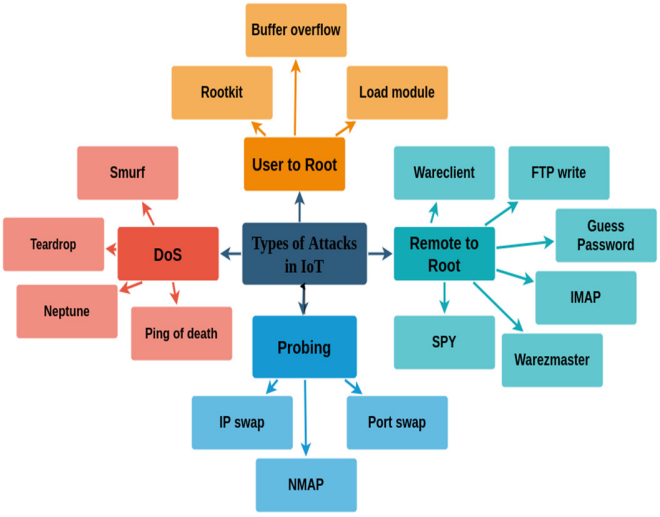

Fig. 3. Attack categories of the IoT ecosystem.

Our data is converted into feature vectors. The dataset is then split into 80% for training and 20% for testing sets (Tables I-II). For the learning algorithm, the training data set was utilized and our final model was deployed using a boosting technique. Figure 4 shows the data distribution in testing and training subsets.

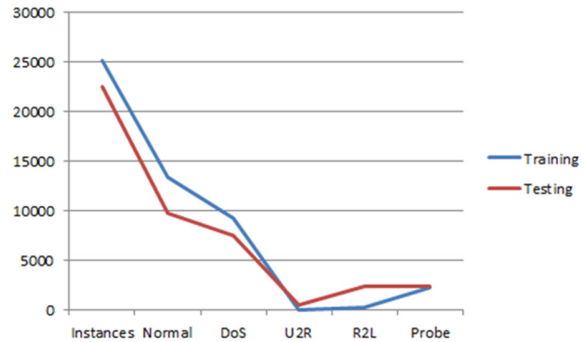

Fig. 4. NSL-KDD dataset distribution.

C. Algorithm

The algorithmic steps are mentioned below.

- Load the NSL-KDD data set.
- Apply the pre-processing technique.
- Divide into 80-20 ratios of testing and training datasets.
- Select feature selection vectors.
- The training dataset is given to the classifiers.
- The test data set is fed to the three selected classifiers for classification.
- Calculate accuracy, specificity, FPR, and TPR.

D. Classifiers and Training

For the model training, RF supervised ML algorithm was selected. The algorithms which combine DT with ensemble learning have several advantages, such as their need of only a few input parameters and their resistance to overfitting. The number of tree parameters is set to be 500. When the number of branches increases, the variance would be decreased without ensuing in bias. RF has changed into applied to traffic data sets, which include in-network traffics misuse detection and Command and Control (C&C) IoT attack detection from traffic flow-base [33].

E. Performance Metrics

In the suggested framework, four performance metrics were considered: Accuracy (A), Training Time (TT), which is the total time to train a classifier, Specificity (S), and Prediction Time (PT), which is the total time which an algorithm takes to predict all the data. TP (true positive) represents the correct identification of an attack, FP (false positive) represents the incorrect identified attacks, TN (true negative) represents the correctly identified normal connections, and FN (false negative) represents the number of attacks that were not correctly identified [34]:

$$\text{Accuracy} = A$$

$$\text{True Positive} = \Theta$$

$$\text{False Positive} = \xi$$

$$\text{True Negative} = \omega$$

$$\text{False Negative} = \Pi$$

Accuracy shows how accurately the algorithm can detect the normal and attack connections:

$$A = \frac{\Theta + \omega}{\Theta + \xi + \omega + \Pi} \quad (1)$$

Specificity is used for measuring the negatives which are correctly identified:

$$S = \frac{\omega}{\xi + \omega} \quad (2)$$

Roc gives a graphical representation that compiles the review of a classifier's overall thresholds on a diagnostic criterion. That is created on mapping the True Positive Rate (TPR) against the False Positive Rate (FPR) as the use of the

threshold is different for selecting algorithms for a provided class:

$$FPR = \frac{\xi}{\xi + \omega} \quad (3)$$

$$TPR = \frac{\Theta}{\xi + \Theta} \quad (4)$$

A threshold is the expected value for all the predicted classes. The ROC curve can be drawn using binary classes. The values of the TPR and FPR range from 0 to 1.

F. Experimental Setup

The experiments were conducted on a Lenovo Thinkpad system with Ubuntu 20.04 operating system, 4500U Processor, 8GB memory, integrated AMD (attached NVIDIA) graphic card which was used for training the dataset. During data preprocessing, cleaning, and feature selection, Numpy and Pandas libraries were used.

G. Result Analysis

As mentioned above, three ML algorithms were applied to the NSL-KDD dataset, namely RF, GDBT, and SVM. From the cross-validation, RF has performed best in terms of testing and training accuracy. The results show that the RF obtained the highest accuracy on fog layer which is 85.34%. The obtained accuracy of SVM and GDBT was 32.38% and 78.01% respectively, as shown in Table III. In terms of specificity, GDBT algorithm performed best with 97.02%. The specificity achieved by SVM and RF was 2.02% and 95.09% respectively. Table III shows the result of the performance evaluation of the mentioned algorithms including *A*, *TT*, *PT*, and *S*.

TABLE III. RESULT ANALYSIS TABLE

Method	*A*	*S*	*TT*	*PT*
SVM, RF	32.38	2.02	10.87	1.056
GDBT	78.01	97.02	7.78	1.6
RF	85.34	95.09	6.10	1.345

VI. CONCLUSION

Through the obtained results, it can be confirmed that supervised ML can be used to analyze traffic data and accurately expose the data that are maliciously traveling over IoT devices. To identify that traffic accurately, NSL KDD dataset is critically evaluated by making use of ML techniques. This dataset is used for the comparison of the given framework by employing functions such as selection and classification. Overall, the RF algorithm provided the best accuracy of 85.34% on the fog layer in comparison with the other two learning algorithms. In the future, it is planned to analyze different IoT devices, explore further technologies and, testing with different data of IoT devices infected by malware and cyber-attacks.

ACKNOWLEDGEMENT

The authors would like to thank the Department of Computer Science and Information technology, NED University of Engineer and Technology for giving permission for this research project.

REFERENCES

[1] S. Mendhurwar and R. Mishra, "Integration of social and IoT technologies: architectural framework for digital transformation and cyber security challenges," *Enterprise Information Systems*, vol. 15, no. 4, pp. 565–584, Apr. 2021, https://doi.org/10.1080/17517575.2019.1600041.

[2] Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities," *Cities*, vol. 89, pp. 80–91, Jun. 2019, https://doi.org/10.1016/j.cities.2019.01.032.

[3] K. K. Mohbey, "An Efficient Framework for Smart City Using Big Data Technologies and Internet of Things," in *Progress in Advanced Computing and Intelligent Engineering*, Singapore, 2019, pp. 319–328, https://doi.org/10.1007/978-981-13-0224-4_29.

[4] N. T. Archibald, "Cybersecurity and Critical Infrastructure: An Analysis of Securitization Theory," *Undergraduate Journal of Politics, Policy and Society*, vol. 3, no. 1, pp. 39–54, 2020.

[5] A. Elsaeidy, I. Elgendi, K. S. Munasinghe, D. Sharma, and A. Jamalipour, "A smart city cyber security platform for narrowband networks," in *27th International Telecommunication Networks and Applications Conference*, Melbourne, VIC, Australia, Nov. 2017, pp. 1–6, https://doi.org/10.1109/ATNAC.2017.8215388.

[6] M. Ghobaei-Arani, A. Souri, and A. A. Rahmanian, "Resource Management Approaches in Fog Computing: a Comprehensive Review," *Journal of Grid Computing*, vol. 18, no. 1, pp. 1–42, Mar. 2020, https://doi.org/10.1007/s10723-019-09491-1.

[7] H. Wang, Z. Wang, and J. Domingo-Ferrer, "Anonymous and secure aggregation scheme in fog-based public cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 712–719, Jan. 2018, https://doi.org/10.1016/j.future.2017.02.032.

[8] D. Li, L. Deng, W. Liu, and Q. Su, "Improving communication precision of IoT through behavior-based learning in smart city environment," *Future Generation Computer Systems*, vol. 108, pp. 512–520, Jul. 2020, https://doi.org/10.1016/j.future.2020.02.053.

[9] A. Churcher *et al.*, "An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks," *Sensors*, vol. 21, no. 2, Jan. 2021, Art. no. 446, https://doi.org/10.3390/s21020446.

[10] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, May 2018, https://doi.org/10.1016/j.future.2017.08.043.

[11] B. K. Mohanta, U. Satapathy, and D. Jena, "Addressing Security and Computation Challenges in IoT Using Machine Learning," in *Advances in Distributed Computing and Machine Learning*, Singapore, Asia, 2021, pp. 67–74, https://doi.org/10.1007/978-981-15-4218-3_7.

[12] J. Li and B. Sun, "A Network Attack Detection Method Using SDA and Deep Neural Network Based on Internet of Things," *International Journal of Wireless Information Networks*, vol. 27, no. 2, pp. 209–214, Jun. 2020, https://doi.org/10.1007/s10776-019-00462-7.

[13] N. Sahar, R. Mishra, and S. Kalam, "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*, Singapore, 2021, pp. 39–50, https://doi.org/10.1007/978-981-15-8377-3_4.

[14] S. Kavitha, U. Maheswari, and R. Venkatesh, "Network Anomaly Detection for NSL-KDD Dataset Using Deep Learning," *Information Technology in Industry*, vol. 9, no. 2, pp. 821–827, Mar. 2021, https://doi.org/10.17762/itii.v9i2.419.

[15] H. Neuschmied, M. Winter, K. Hofer-Schmitz, and B. Stojanovic, "Two Stage Anomaly Detection for Network Intrusion Detection," in *7th International Conference on Information Systems Security and Privacy*, Vienna, Austria, Feb. 2021, pp. 450–457.

[16] S. Pande, A. Khamparia, D. Gupta, and D. N. H. Thanh, "DDOS Detection Using Machine Learning Technique," in *Recent Studies on Computational Intelligence: Doctoral Symposium on Computational*

Intelligence, A. Khanna, A. K. Singh, and A. Swaroop, Eds. Singapore, Asia: Springer, 2021, pp. 59–68.

[17] A. T. Siahmarzkooh, J. Karimpour, and S. Lotfi, "A Cluster-based Approach Towards Detecting and Modeling Network Dictionary Attacks," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1227–1234, Dec. 2016, https://doi.org/10.48084/etasr.937.

[18] A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng, "Fog Computing for the Internet of Things: Security and Privacy Issues," *IEEE Internet Computing*, vol. 21, no. 2, pp. 34–42, Mar. 2017, https://doi.org/10.1109/MIC.2017.37.

[19] I. Kotenko, I. Saenko, A. Kushnerevich, and A. Branitskiy, "Attack Detection in IoT Critical Infrastructures: A Machine Learning and Big Data Processing Approach," in *27th Euromicro International Conference on Parallel, Distributed and Network-Based Processing*, Pavia, Italy, Feb. 2019, pp. 340–347, https://doi.org/10.1109/EMPDP.2019.8671571.

[20] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. M. A. Hashem, "Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches," *Internet of Things*, vol. 7, Sep. 2019, Art. no. 100059, https://doi.org/10.1016/j.iot.2019.100059.

[21] A. Haldorai, A. Ramu, and M. Suriya, "Organization Internet of Things (IoTs): Supervised, Unsupervised, and Reinforcement Learning," in *Business Intelligence for Enterprise Internet of Things*, A. Haldorai, A. Ramu, and S. A. R. Khan, Eds. Cambridge, UK: Springer, 2020, pp. 27–53.

[22] S. Rathore and J. H. Park, "Semi-supervised learning based distributed attack detection framework for IoT," *Applied Soft Computing*, vol. 72, pp. 79–89, Nov. 2018, https://doi.org/10.1016/j.asoc.2018.05.049.

[23] M. Lopez-Martin, B. Carro, and A. Sanchez-Esguevillas, "Application of deep reinforcement learning to intrusion detection for supervised problems," *Expert Systems with Applications*, vol. 141, Mar. 2020, Art. no. 112963, https://doi.org/10.1016/j.eswa.2019.112963.

[24] K. Yang, J. Ren, Y. Zhu, and W. Zhang, "Active Learning for Wireless IoT Intrusion Detection," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 19–25, Dec. 2018, https://doi.org/10.1109/MWC.2017.1800079.

[25] B. S. Bhati and C. S. Rai, "Analysis of Support Vector Machine-based Intrusion Detection Techniques," *Arabian Journal for Science and Engineering*, vol. 45, no. 4, pp. 2371–2383, Apr. 2020, https://doi.org/10.1007/s13369-019-03970-z.

[26] M. M. N. Aboelwafa, K. G. Seddik, M. H. Eldefrawy, Y. Gadallah, and M. Gidlund, "A Machine-Learning-Based Technique for False Data Injection Attacks Detection in Industrial IoT," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8462–8471, Sep. 2020, https://doi.org/10.1109/JIOT.2020.2991693.

[27] L. Liu, J. Yang, and W. Meng, "Detecting malicious nodes via gradient descent and support vector machine in Internet of Things," *Computers & Electrical Engineering*, vol. 77, pp. 339–353, Jul. 2019, https://doi.org/10.1016/j.compeleceng.2019.06.013.

[28] M. B. Farukee, M. S. Z. Shabit, Md. R. Haque, and A. H. M. S. Sattar, "DDoS Attack Detection in IoT Networks Using Deep Learning Models Combined with Random Forest as Feature Selector," in *Advances in Cyber Security*, Singapore, Asia, 2021, pp. 118–134, https://doi.org/10.1007/978-981-33-6835-4_8.

[29] P. A. A. Resende and A. C. Drummond, "A Survey of Random Forest Based Methods for Intrusion Detection Systems," *ACM Computing Surveys*, vol. 51, no. 3, pp. 48:1-48:36, May 2018, https://doi.org/10.1145/3178582.

[30] R. Kozik, M. Choras, M. Ficco, and F. Palmieri, "A scalable distributed machine learning approach for attack detection in edge computing environments," *Journal of Parallel and Distributed Computing*, vol. 119, pp. 18–26, Sep. 2018, https://doi.org/10.1016/j.jpdc.2018.03.006.

[31] G. Meena and R. R. Choudhary, "A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA," in *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, Jaipur, India, Jul. 2017, pp. 553–558, https://doi.org/10.1109/COMPTELIX.2017.8004032.

[32] J. Liu, B. Kantarci, and C. Adams, "Machine learning-driven intrusion detection for Contiki-NG-based IoT networks exposed to NSL-KDD dataset," in *2nd ACM Workshop on Wireless Security and Machine Learning*, New York, NY, USA, Jul. 2020, pp. 25–30, https://doi.org/10.1145/3395352.3402621.

[33] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, "Rule Generation for Signature Based Detection Systems of Cyber Attacks in IoT Environments," *Bulletin of Networking, Computing, Systems, and Software*, vol. 8, no. 2, pp. 93–97, Jul. 2019.

[34] M. Aamir and S. M. A. Zaidi, "DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation," *International Journal of Information Security*, vol. 18, no. 6, pp. 761–785, Dec. 2019, https://doi.org/10.1007/s10207-019-00434-1.

[35] Y. L. Ng, X. Jiang, Y. Zhang, S. B. Shin, and R. Ning, "Automated Activity Recognition with Gait Positions Using Machine Learning Algorithms," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4554–4560, Aug. 2019, https://doi.org/10.48084/etasr.2952.

[36] Z. A. Shaikh, "Keyword Detection Techniques: A Comprehensive Study," *Engineering, Technology & Applied Science Research*, vol. 8, no. 1, pp. 2590–2594, Feb. 2018, https://doi.org/10.48084/etasr.1813.

[37] N. F. Syed, Z. Baig, A. Ibrahim, and C. Valli, "Denial of service attack detection through machine learning for the IoT," *Journal of Information and Telecommunication*, vol. 4, no. 4, pp. 482–503, Oct. 2020, https://doi.org/10.1080/24751839.2020.1767484.