SECURITY ESTABLISHMENT IN INTERNET OF THINGS TO MANAGE TRAFFIC USING INTEGRATED LEARNING APPROACHES

*S.Nanadhini** [1], *Thilagavathi* [2]

ABSTRACT

The Internet of Things (IoT) has seen an extraordinary growth rate recently; meanwhile, cybercrime activities have also gained unwanted attention, posing security threats. This is proved by the number of communication media and IoT devices being attacked by cybercrime. The IoT users will face severe losses in finances and service interruption when the attacks are not quickly detected and rectified. Cyber-attacks will impose an identity protection threat. Real-time intrusion detection is essential to make services based on the Internet of Things reliable, profitable and secure. This study projects a contemporary system for intrusion detection in IoT devices based on deep learning (DL). This paper introduces an intelligent system with deep network architecture to identify the malicious traffic that might attack IoT devices. To decrease the complexities in deployment, The proposed Deep Learning-based Intrusion Detection (DLID)-based Xtreme Gradient Boosting (XGB) system has been built as a model independent of the communication protocol. During the analysis of the proposed system's performance, reliable outcome is achieved in both real-time and simulated environments. Our model's accuracy of 94.75% on average is completed by detecting anomalies such as distributed DoS, black hole detection, wormhole attacks, sinkholes, and opportunity services. Precision value of 94.71%, F1-score of 94.82% and recall value of 95.47% is achieved on average by the system proposed for intrusion detection. IDS systems based on innovative deep learning provide an average detection rate of 93.2%, considered acceptable for IoT network security.

Keywords: IoT, network security, machine learning, deep leanring, prediction

Department of Computer Science and engineering[1],
Karpagam Academy of Higher Education, Coimbatore, India[1]
nandhini.sivaraj@kahedu.edu.in
NGM College, Pollachi[2]
thilagavathibluesky@gmail.com
* Corresponding Author

I. INTRODUCTION

This computing system brings about a new shift in people's lives called The Internet of Things (IoT). In a vast number of applications, IoT devices are used. A few of these applications are as follows: healthcare, smart homes, including transportation, supply-chain management [1], industrial production, Blockchain monitoring and security. The author have carried out research which states that by 2025, devices connected to IoT will reach a whooping number of 30.9 billion. A previous study says that 13.8 billion [2] IoT devices were present in 2021. This shows that this industry achieves 55% of the growth rate. So, because of its rapid growth, the business potential has attracted the attention of corporate people, researchers, devices dependent on IoT, innovators, entrepreneurs, business people, and lastly, cyber-criminals [3]. In the market, there is a high demand for the services provided by devices connected to IoT. So, potential investors are showing great interest in this new field. To make people's lives so much easier, entrepreneurs and innovators are developing various attractive and promising services based on IoT devices. However, the same interest for a better life has enabled criminals to attack vulnerable, weak devices connected to IoT cyber-attacks with fewer security features. Cyber-attackers are taking this opportunity to practice their malicious activities in attractive domains like finance and business. This is the primary reason for the growing number of attacks on devices connected to IoT [4].

The security of IoT devices is studied by [5], who express that when manufacturing IoT devices, very little attention is paid to cybersecurity or even no attention is paid. Research conducted and proposes that many IoT devices present in the field currently in different services are fully operational but are vulnerable to attacks from cyber criminals [6]. Attacks are not confined to IoT devices. Devices with IoT service are not stand-alone devices. They are connected to various other systems and appliances, which makes them an attack vector, allowing the attackers to exploit the devices connected to

them by gaining their access. In 2016, primary internet-dependent services like Netflix, Amazon and Twitter faced devastating effects of internet outages due to imaging IoT devices connected through Dyn coming under the cyber-attack [7]. Medical IoT, or MioT, is the fastest-rising sector based on IoT service usage. Since healthcare is a sensitive field and needs strict privacy, an attack on MIoTs will be devastating. The trends in cyber-attacks in recent times show that the MIoT device intrusion rate is higher than ever, and it is snowballing at a high pace.

Security vulnerabilities in IoT devices are the primary shortcoming faced by the Internet society. IoT devices with vulnerabilities can be exploited more ways than we think [8]. Systems connected to IoT devices are constrained computationally. Other than cyber-physical attacks (out of the scope of this research), every other cyber-attack is made possible using network connectivity, which is not within this scope. It is impractical to secure vulnerable IoT environments using conventional methods for cyber-security.

Furthermore, there are several ways the vulnerabilities of IoT devices can be exploited, so the traditional or rule-based security methodology needs to fit the security requirements of today's internet world. Vast volumes of data are constantly produced in IoT environments. It is challenging to detect anomalous data from the sea of legitimate data. In cases like these, approaches based on Deep Learning (DL) are appropriate. Analyzing and classifying massive data, discovering patterns among them, and finding relations based on the properties defined are challenging tasks, and deep neural networks are highly capable of doing this analysis. Hence, it is used in our proposed system.

In IoT security and intrusion detection based on anomaly, a promising security solution is provided using Deep Neural Networks (DNN) applications. Even though the IoT environment tends to produce large volumes of data, a similar pattern is formed by data due to its shared nature. Any data that does not follow the natural pattern will be considered abnormal. A network that is adequately trained can be a potential solution for identifying and classifying anomalous data [9]. A system for detecting intuition from real-time anomalous data in an IoT environment based on a proposed model with the help of the hypothesis provided above. This research is being put forward to build an independent IoT intrusion detection system with a communication protocol through a deep neural network to provide a secure and safe IoT environment.

In this research, we proposed a (DLID) Deep Intrusion-Detection system, a IDS. The main advantage of this system is that it does not need network structure virtualization, system attributes or any communication protocol modification. When DLID is connected to a network, it will detect the standard data flow and anomalous signal flow. Intrusions are detected by finding the data flow signals that deviate from the consistent signal pattern. This paper provides the following core contributions:

- ❖ Design and optimization of DLID-based Xtreme Gradient Boosting for intrusion detection in a network connected to an IoT environment.
- ❖ Structuring of components embedded in the system for intrusion detection.
- ❖ They are identifying excellent features for datasets to train the IoT network to detect intrusion effectively.
- ❖ An intrusion model with commonly occurring attacks of five numbers was used for the DLID-based intrusion detection system's thorough analysis.
- ❖ An average rate of 94.7% intrusion detection is achieved in a network.

The work is organized as: section 2 describes the literature work. The methodology is provided in section 3. The numerical results are given in section 4 and conclusion in section 5.

II. RELATED WORKS

IoT devices and their services have made our day-to-day lives so much easier. Most of the sectors and their access have been revolutionized by their application. Even though IoT devices achieve a remarkable growth rate, their adoption by people from all industries could be better due to cybersecurity concerns. It poses a significant barrier to the adoption pace of IoT technologies. When an IoT device is attacked successfully, other devices connected to it will be

prone to be explored by cyber-criminals, so it causes a cascading effect on the devices connected to IoT. At the application level, IoT devices are extensively used; they are commonly used to make human personal aspects easier. This is one of the primary reasons cyber-criminals target attacks on IoT devices [11].

Furthermore, big manufacturing units control the units' activities through IoT device automation. If a breach in security is to happen, then the whole system and its access will be in the hands of the attacker. During situations like this, a harmless smartphone or watch connected to IoT will become vulnerable to such unwarranted attacks. These features make cybercriminals interested in attempting to attack extensive facilities connected to IoT. In order to reduce the number of cyber-attacks on IoT devices, we should develop an intrusion detection system based on Deep Learning.

In their review paper, the author proposed different methods and their effectiveness in detecting and eliminating intrusions in IoT environments. The authors exhibit an excellent performance rate for detecting intrusion in IoT network built the SHAP model by using a deep learning technique. They employed the network's Shapley additive explanations (SHAP) and exhibited improved results. The primary feature of this system is that the Deep Learning-based IDS's decision-making is analyzed logically to understand the working principle. The SHAP framework exhibits a promising 99.15% and 98.83% accuracy and F1 score, respectively. The field of interest for the SHAP framework is (IoV) Internet of Vehicles. It is an IoT sub-branch. The author in [12] stated that intrusion detection systems based on ML work exceptionally in environments like edge and fog computing. The proposed intrusion detection system works optimally for any IoT network without constraints to the connected devices. In the meantime, an accuracy of 99.15% is attained by A. Oseni et al., which brings some doubt about excessive data, and it has to be tested with various datasets with diverse data to address the suspicion. Our system is more robust and reliable when the above framework is compared with the proposed IDS.

A deep Learning-based system named DeepIIoT was established. It intends to provide proper security to IoT devices connected to the industry; it succeeded in its mission by exhibiting an accuracy of 99% when executed with the collected testing dataset. Their area of interest is primarily power grids, facilities for water treatment, and nuclear reactors based on heavy industrial IoT devices. A dataset named WUSTL-IIOT-2021 is employed for training and testing the framework [13]. The deep learning-based IDS is proposed to cover all IoT devices and to provide better security at every level. However, network-independent IoT service is provided by the novel intermediate communication system. Unlike other sector-specific applications, the IDS system and its application offer services for various sectors. Furthermore, the DeepIIoT system is a detection system based on a rule and depends on the signature for processing. One of the challenges faced is recurrent updating, which is needed for every database use. There are no such barriers in the system we proposed. In the framework presented by V. Ravi et al., an ensemble meta-classifier is employed in the intelligent network intrusion detection system they developed; it uses a fusion approach with recurrent deep learning features. The framework mentioned above exhibits an average accuracy of 99% [14]. Even though the above method provides a higher accuracy rate, its network architecture is complex and challenging to deploy, whereas the proposed method overtakes simple network architecture. So, it gives a fresh perspective on the IoT domain and its future enhancements. The author presented a methodology for intrusion detection using mining-based ANN. Intrusion detection in communication networks demonstrates substantial improvement. ANNs are employed in the proposed approach. However, network architecture is particularly premeditated for identifying IDS in IoT networks. Meaning it is domain-specific.

DL techniques in IoT security is expressed in IoT privacy can also use this application and its uses in future work, as they emphasized [15]. Developing rule-based intrusion detection systems specific to certain networks is easy because they will perform well only within the particular network. H G. An et al. developed a sinkhole attack intrusion detection system, a knowledge-based network dependent on

a specific rule. This system provides an 81.63% attack detection ratio, which is 7% better performance than; an approach named INTI is presented. An open course networks for rule-based intrusion detection systems and their variations in dynamic analysis, discussed the importance of using suitable IDS for different configurations and networks [15]. The motto of the methodology proposed and the above paper's quantitative analysis fall perfectly together; this is needed for intrusion detection systems in a network applied with IoT, independent of the protocol used.

III. METHODOLOGY

3.1. Dataset

For this experiment, 25,000 instances were arranged, which constitutes one dataset. In the provided dataset, 82:18 is the instance ratio of Regular (R_D) to Malicious (M_D). It can be explained as, out of 25,000 instances, malicious data accounts for 18% and regular data accounts for 82%. In the prepared dataset, no pattern was formed intentionally, and M_D. and R_D are mixed randomly. In the ratio of 70:15:15, the data in the dataset were separated into training, testing, and validation data. During the training phase, the validation and training datasets have been utilized. Only during the phase of experimental evaluation is the testing dataset put into use; it is kept undisturbed during the validation and training phase. In systems for automatic intrusion detection, it is essential to perform appropriate feature reduction and feature selection for enhanced detection. So, for this experiment, An ML algorithm-based intrusion detection system determines the influential features for better outcomes.

3.2. Method

The DLID system proposed and represented in Fig 1. This is an automatic system with cutting-edge capabilities. The DLID learned the host IoT network and the data they generated, and the intrusion was detected as soon as it received sufficient training. The proposed IDS's dynamic connector will initiate the link between the emulated network, and the IoT network is asked for the request fulfilment. Using an interface module, the network classifier and feature extractor will communicate in the simulated network. The network packet's features are extracted through a feature extractor used as the proposed deep neural network dataset. With the help of the classifier Updater module, the recently identified features are updated regularly. That's why the IDS system proposed is said to be dynamic. Whenever an intrusion is found to have happened, the network classifier will pass it to mitigation stage. So, reducing the intrusion impact of this phase. Quantitative and qualitative features are presented. Receivers and senders will be in the network connected to IoT. Devices connected to IoT are designed to communicate in the least complex way possible because these devices are constrained to use resources. The receiver's and senders' transmission rates are kept the same to eliminate the need to install the extended buffer memory. DoS attack can be defined as a substantial variation in the distribution and transmission rate.

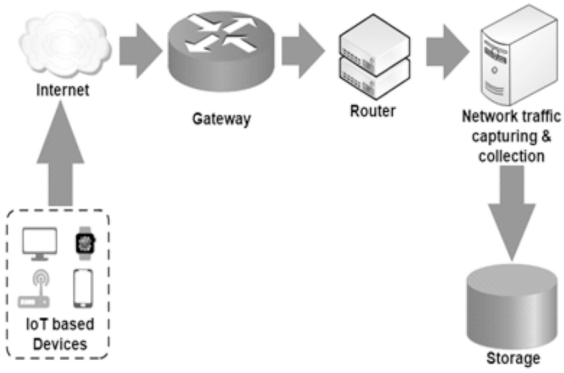

Figure 1 : Umbalanced network traffic

IoT network is safe when the reception and transmission ratio is near unity. Moreover, the ratio deviates substantially from unity. In that case, it indicates that the receiver is not getting data from intended source and the sender is not transferring it to the intended source. This indicates spoofing. There are some unlisted intermediate devices by which the security has been breached, and an external source manipulates the receiver and the sender. The communication protocol determines the communication cycle status, the following phase, and the transmission mode. A DLID is trained to communicate under various transmission modes. When some device behaves unfamiliarly in a particular transmission mode, the network is found to have intruded. In

an IoT network, qualitative results will be the outcome of quantitative analysis, which will help the system confirm intrusion in an IoT device. Quantitative analysis uses the features presented to produce qualitative results. If a device is under attack, the DLID will use the Communication Module's cache memory to store these features temporarily. It will use them to spot the possible intrusion successfully. This is done by calculating the probability distribution with the help of the following Eq. (1).

$$P_d = P(f_0), P(f_1), P(f_2), \dots, P(f_7) \quad (1)$$

Here is Eq. (1), $P(f_i)$ specifies probability of every data packet identified in i^{th} feature. If the mean of the $P(f_i)$ is above 50%, then the data is malignant/benign.

ML-based algorithm uses gradient-boosted model to boost performance and speed. Compared to other gradient-boosting methods, XGB shows outstanding speed. The blend of earlier models' ideas was used to form the new models to achieve ultimate outcomes. The proposed model's performance will be improved, and loss will be brought down by employing this gradient descent method. When ML-based problems are analyzed, the above process is appreciated by data scientists as it provides an end-to-end tree-boosting technique. Two primary components are embedded within XGB: regularization and training loss, θ denoting the ideal training data " and XGBoost.

$$O(\theta) = L(\theta) + \Omega \quad (2)$$

Here, the training loss function is represented by L , the metric used to evaluate model's performance in predicting the training samples. Mean Squared Error (MSE) is represented as a training loss function:

$$L(\theta) = \sum_i (y_i - \hat{y}_i)^2 \quad (3)$$

In logistic regression, the frequently employed loss function is the logistic loss function:

$$L(\theta) = \sum_i [y_i \ln(1 + e^{-\hat{y}_i}) + (1 - y_i) \ln(1 + e^{-\hat{y}_i}) \quad (4)$$

The regularization process controls the model's complexity; this factor helps by eliminating the overfitting issue, and it is given as follows:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} ||w||^2 \quad (5)$$

IV. NUMERICAL RESULTS

This study uses Python with the TensorFlow library to implement the proposed DLID system for IoT networks. Cooja Network Simulator (CNS) is employed here. The Contiki Operating System (COS) is used for experimentation. The introduced intrusion detection system's experimentation and performance were evaluated using metrics derived from cutting-edge ML techniques. The researchers found that precision, accuracy, F1-score, and recall are the evaluation metrics broadly employed based on ML. Eq. (6) to Eq. (10) define these metrics in the order of accuracy, recall and F1-score. False Negative (FN), True Negative (TN), False Positive (FP) and True Positive (TP) are used to measure the evaluation metrics. The confusion matrix analysis present in the following subsection contains these values. The proposed DLID system's intrusion Detection Rate (IDR) performance is evaluated, and Equation (17) presents the same. Here, Detected Intrusion is given as DI, and Total Intrusion is given as TI.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad (9)$$

$$IDR = \frac{DI}{TI} * 100\% \quad (10)$$

Table 1 lists overall performance based on intrusion type. The DLID system's comprehensive performance analysis is done by experimenting with random attacks in a live environment. Owing to confidentiality and security issues, we DLID not allow unknown attackers to exploit our system by opening it. Nevertheless, the DLID system's performance is evaluated using a realistic intrusion model; for this experimentation, our system is monitored round the clock (24h). Many random attacks happened during this time. Detected intrusions were thus made and logged for evaluation and record.

Table 1 Performance comparison

Attack	Acc	Pre	Re	F1	IDR
Blackhole	94	94.6	96.1	94.2	98.5
DoS	94.9	94.8	95.3	94.6	95.4
OS	96.2	96.1	96.3	95	91
Sink hole	94	94.3	94.2	93.8	92.8
Worm hole	95.6	95.6	94.5	94.3	92.3
Average	**94.7**	**94.7**	**93.8**	**94.4**	**93.2**

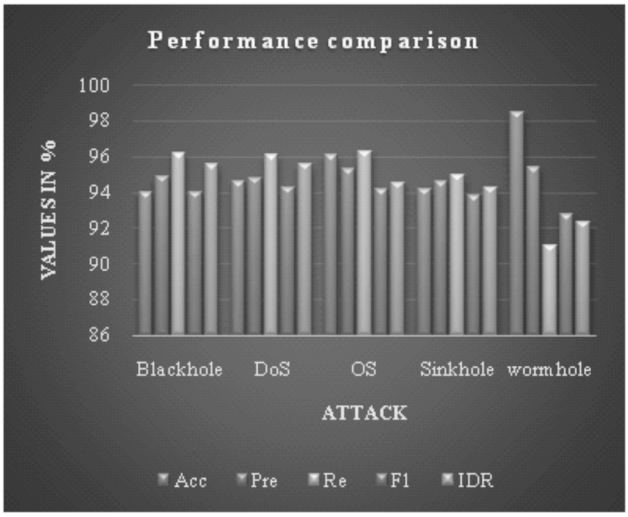

Figure 2 : Performance comparison

Fig. 2 demonstrates the system's overall performance. The deep learning-based IDS has exhibited the best outcome in identifying (OSA) Opportunistic Service Attacks. Yet, it does not mean that DLID needs to perform better in other systems. The proposed system for all intrusion detection exhibits 94.74% accuracy on average, 94.172% precision, 95.824% recall, and 95.472% F1 score. For all five attack types, the intrusion detection rate is observed. DLID protects against most experimenting attacks, and 95.11% of the average IDS is kept. Fig 2 illustrates that borderline variation is observed in average accuracy, precision, recall, and F1-score. Evaluation metrics' average value range is 94% to 95%, presented in yellow. This indicates that the DLID system detects intrusion automatically in practical applications on IoT networks in a stable, suitable and accurate way, hence better protecting the network's devices.

V. CONCLUSION

With the present boom in the IoT sector, smart devices connected to IoT communicate with one another and themselves to perform required tasks. This scenario is backed by the 30.9 billion active IoT devices. The vast 65% development in the IoT sector makes business people, entrepreneurs, educationalists, and cybercriminals gravitate towards IoT. Owing to the computing environment's resource-constrained feature, IoT device manufacturers need to pay attention to security while concentrating on quality of service. Attacks made on IoT do not end on the same device but will move to all other devices connected to it, jeopardizing the entire network. That is why it attracts cybercriminals more than ever in the internet-dependent environment. This influenced us to develop a system for intrusion detection as soon as attackers make an attempt, which we found to be much more essential. The novel deep learning-based IDS detects frequent and commonly attempted attacks with 94.74% average accuracy. A deep learning-based based DLID-XGB is trained for a specific network performs fine over the heterogeneous environment. The proposed system collects datasets from various from different networks. This novel method is a fresh wave of air for IoT device security as it provides a scalable and robust intrusion detection system and automatically detects intrusion; it also can learn from other IoT networks. A platform-independent framework that can be developed for IoT networks from the DL-based IDS will be explored for future scope. It is anticipated that a proposed framework's lightweight version will be much more effective and efficient in detecting intrusion.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.

[2] Reddy, Y. Ramadevi, and K. V. N. Sunitha, "Effective discriminant function for intrusion detection using SVM," in Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICAC), Sep. 2016, pp. 1148–1153.

[3] Ingre and A. Yadav, "Performance analysis of NSL-KDD dataset using ANN," in Proc. Int. Conf. Signal Process. Commun. Eng. Syst., Jan. 2015, pp. 92–96.

[4] Farnaaz and M. A. Jabbar, "Random forest modelling for network intrusion detection system," Procedia Comput. Sci., vol. 89, pp. 213–217, Jan. 2016.

[5] Khan and N. Jain, "A survey on intrusion detection systems and classification techniques," Int. J. Sci. Res. Sci., Eng. Technol., vol. 2, no. 5, pp. 202–208, 2016

[6] Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software-defined networking," Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM), Oct. 2016, pp. 258–263.

[7] Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He, "Fuzziness based semi-supervised learning approach for an intrusion detection system," Inf. Sci., vol. 378, pp. 484–497, Feb. 2017.

[8] Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He, "Fuzziness based semi-supervised learning approach for an intrusion detection system," Inf. Sci., vol. 378, pp. 484–497, Feb. 20

[9] Chang, W. Li, and Z. Yang, "Network intrusion detection based on random forest and support vector machine," Proc. IEEE Int. Conf. Comput. Sci. Eng./IEEE Int. Conf. Embedded Ubiquitous Comput., Jul. 2017, pp. 635–638

[10] Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring: A survey," Submitted to IEEE Trans. Neural Netw. Learn. Syst., 2016. [Online]. Available: http://arxiv.org/abs/1612.07640

[11] Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," J. Mach. Learn. Res., vol. 11, pp. 3371–3408, 2010.

[12] In Proc, you, Y. Li, Y. Wang, J. Zhang, and Y. Yang, "A deep learning-based RNNs model for automatic security audit of short messages," in Proc. 16th Int. Symp. Commun. Inf. Technol., Qingdao, China, Sep. 2016, pp. 225–229.

[13] Sherubha, "Graph Based Event Measurement for Analyzing Distributed Anomalies in Sensor Networks", Sådhanå(Springer), 45:212, https://doi.org/10.1007/s12046-020-01451-w

[14] Sherubha, "An Efficient Network Threat Detection and Classification Method using ANP-MVPS Algorithm in Wireless Sensor Networks", International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8 Issue-11, September 2019

[15] Sherubha, "An Efficient Intrusion Detection and Authentication Mechanism for Detecting Clone Attack in Wireless Sensor Networks", Journal of Advanced Research in Dynamical and Control Systems (JARDCS), Volume 11, issue 5, Pg No. 55-68