An Intelligent Approach to Increase the Performance of Threat Detection in IoT

[1]R Tamilkodi,[2] V Bala Sankar, [3*]Nersu Pavankumar, [4]Potnuru Hemanth kumar, [5]Uggu Veera Gani Durga, [6]Bokka Durga Pravallika

[123456]Department of Computer Science & Engineering (AIML&CS) Godavari Institute of Engineering & Technology, Rajahmundry, Andhra Pradesh, India

[1]tamil@giet.ac.in,[2]balasankar@giet.ac.in,
[3*]20551a4638.pavankumar@gmail.com,
[4]20551a4643.hemanthkumar@gmail.com,
[5]21555acsc4605.gani@gmail.com,[6]20551a4606.pravallika@gmail.com

Abstract. The ubiquitous use of IoT (Internet of Things) devices is on the increase. In order for an Internet of Things system to function, it includes all of the necessary hardware, software, networks, sensors, and other parts. The developers of these sensors and devices, however, often omitted details about their minimal resource needs and a slew of security vulnerabilities. In addition, there are a lot of risks associated with the placement of edge networks for IoT devices. The system's performance might be severely compromised by denial-of-service assaults or unlawful sensor hijacking on sites inside the edge network. Our paper presents a model for training and forecasting DDoS attacks using principal component analysis and machine learning methods. The data's dimensionality was reduced using principal component analysis techniques. Metrics for evaluation included precision, accuracy, F1score, and recall. Important parts of the evaluation metrics mentioned earlier are Metrics such as True- Positive, False- Positive, True -Negative, and False -Negative are utilized to assess the impact of the Fourth Industrial Revolution. We used the Training Time to compare each model's training time, which differs from past research. With the use of the CICIDS 2017 and CICIDS 2018 datasets, we assess the performance of our suggested model. In comparison to similar models, the suggested models outperformed them while requiring much less time to train.

Keywords: Internet of Things, Machine Learning Models, Threat Detection.

1 Introduction

The recent decade has seen a dramatic shift in the industrial landscape Due to the impact of the Fourth Industrial Revolution, there have been significant changes across various sectors. Beyond transforming business and industry, this revolution has also

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_137

revolutionized human existence and the environment [1]. This revolution is built on four main components: the Internet of Things (IoT), cloud computing, RPA, and artificial intelligence (AI). The Internet of Things (IoT) has had far-reaching implications on many parts of our lives, including healthcare, agriculture, disaster relief, and the independence of people with physical limitations [4, 6]. What we call "the Internet of things" (IoT) is really just a system of networked Devices and systems capable of gathering, transmitting, processing, and analyzing data from the physical environment are instrumental components of the Fourth Industrial Revolution. The Internet of Things (IoT) relies on a number of key components. In order to collect data, analyze it in advance, and undertake initial analysis, the edge devices are crucial. They are able to gather, sort, and transmit relevant data to the cloud, making them the principal data collectors. Gateways mediate communication between cloud servers and edge devices, allowing for more efficient data transfer and a more streamlined connection [4,11]. Contrarily, data processing and storage mainframes are cloud servers. They allow for the analysis and aggregation of massive datasets, which in turn allows for decisions to be made in real-time. The ability to extract useful information from datasets is dependent on data analysis tools. In order to find trends, patterns, and outliers, these tools use a variety of methods, such as visualization, statistical analysis, and machine learning algorithms. [7]. Insights like this are helpful for making educated choices. Users are able to communicate with the Internet of Things system via user interfaces, which are often web or mobile apps. The data collected from their devices may be accessed, reviewed, and used to generate alerts and warnings, empowering them with actionable information.

2 Related Works

2.1 An Industry 4.0 system's real-time hybrid machine learning ensemble is developed for anomaly detection.

Because of the difficulty in adequately covering the complexity of an industrial system, errors and abnormalities in real-time systems may be difficult to detect. Industry 4.0, which is made possible by the development of technologies like machine learning and the Internet of Things, may provide answers to these problems [1]. In order to improve anomaly identification, this work proposes a real-time anomaly detection pipeline that combines three machine learning models: autoencoder, one-class support vector machine, and local outlier factor. The models employ a weighted average for aggregation. We used three different air-blowing devices to test the ensemble model, and the F1-scores. When comparing performance measurements, the ensemble model proved to be superior to the individual measures. The two parts that make up this model are the production and operations stages, which are characteristic of every industrial system. This is what makes this model special.

2.2 A DDOS Attack Risk Assessment and Prevention Study

At this point in time, everything is reliant on the Internet, which provides people all over the world with access to information. So, you can't do without the internet. Distributed denial-of-service assaults are very common and expensive in the cyberattack landscape today. The majority of this article focuses on distributed denial-of-service attacks, which restrict access to networks by overwhelming their targets with an excessive amount of malicious information [8]. We went over the various DDoS assaults that online service providers face. The goal of this study is to find different ways to stop these assaults, limit their strategies, and
provide solutions

2.3 A Signature-Based Method for Detecting Botnets (Emotets)

As a consequence of the COVID-19 pandemic, the Internet has become an integral component of contemporary living. As technology is used more often, new challenges emerge. During this crisis, security has become an issue of paramount importance. In comparison to other forms of hacking, Distributed denial of service (DDoS) attacks have become increasingly prevalent more sophisticated as the pandemic has progressed. A botnet provides the platform for the attacker and answers the most crucial question, "What is the source of the DDoS attack?" [12]. The goal of an escalation campaign by a botnet is to target vulnerable systems. Thus, it is necessary to have effective methods for detecting and preventing bot-nets. For a highly malicious botnet to halt networks in their tracks, IoCs are crucial due to the dataset's organization. Research has made use of several malware datasets, the majority of which are out of date. In order to better understand Windows-based botnets, the author used many analytic methodologies to generate a new dataset [14]. This work provides the location of the malicious emotet connection. Furthermore, they showcased the process of calculating IP reputation and using look intrusion detection to detect botnets that rely on IoCs.

2.4 A novel hybrid machine learning approach for the detecting unprecedented distributed denial of service attacks.

With the prevalence DDoS assaults on computer networks growing annually, service availability is of the utmost importance. As a stochastic method, machine learning (ML) does a decent job at detecting DDoS assaults that are already known to exist. But they seldom identify undiscovered dangerous signals. Combining supervised and unsupervised algorithms is a new approach that this work introduces. Prior to clustering, a number of flow-based criteria are used to differentiate between normal and abnormal traffic. The designations for the groups are determined using a statisticallybased categorization method. Using the CICIDS2017 dataset for training and an additional set of attacks from the more current CICDDoS2019 for testing, we assess the suggested strategy inside a massive data processing framework [18]. Our approach outperforms machine learning classification methods in terms of Positive Likelihood Ratio (LR+) [19].

3 Proposed Methodology

They documented a method for identifying and categorizing problems with IoT networks used in gardening. They go over some of the protection and security issues with horticulture-related IoT networks and demonstrate how to use interruption detection frameworks (IDS) to spot both internal and external assaults on a business's PC infrastructure. To use the NSL KDD dataset as an information dataset, it first goes through a comprehensive process of converting representative components to numerical highlights and back again. After the components are separated using head part analysis, the pre-processed information is organized using artificial intelligence methods like support vector machines and direct relapse. Focusing on review boundaries, accuracy, and correctness, we evaluate the presentation of AI calculations.

For optimal algorithm selection and comprehensive testing, there are several approaches to use algorithms

- To determine whether machine learning approaches are effective, we evaluate and contrast five of them.

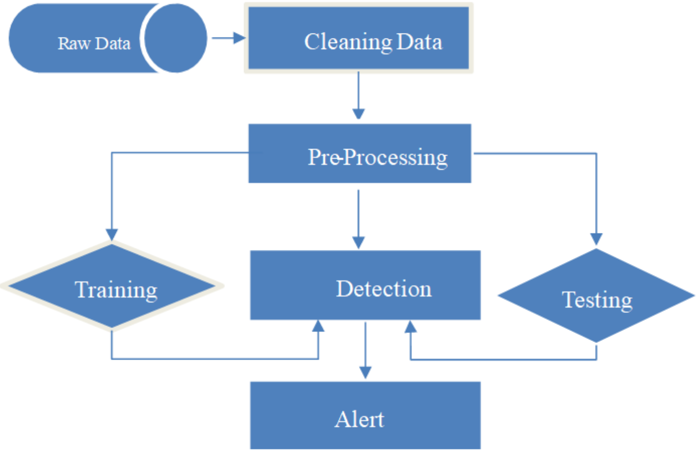


Fig.1. System architecture

3.1MODULES

- Data exploration is the process of entering information into the system.
- The content must be read in order to be handled.
- Train and test sets will be created from the data. For this, we shall use this instrument.

  The following methods will be used for model construction: The analysis included models such as RF, DT, Extra Tree, Naïve Bayes, SVM, and Voting and Stacking Classifier.
- Input from the user: This part takes in the user's input for the forecast.
- Users have the option to register and log in; this module manages both processes.
- The final prediction is shown in the prediction section.

4 Implementation

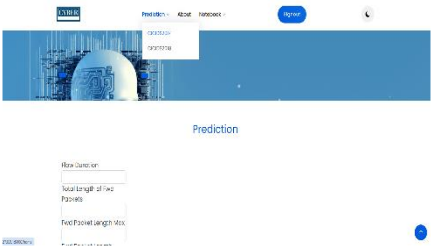
4.1 Voting classifier (RF + AdaBoost):

One kind of ensemble learning is the voting classifier, which takes many base models and uses them all to choose the best one. To forecast specific outcomes, the underlying model may operate autonomously using various algorithms like KNN, Random forests, Regression, and so on.

4.2 Stacking classifier (RF + MLP with LightGBM)

As an ensemble approach, stacking classifiers include feeding the results of many classifiers into a meta-classifier, which is then responsible for making the final classification. For multi-classification problems, the stacking classifier method may be a lifesaver.

5 Results and Discussion



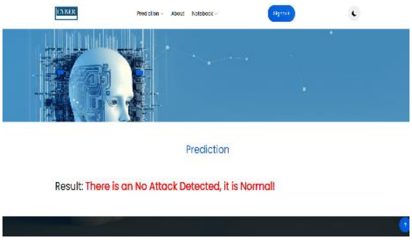
Prediction

Flow Duration

Total Length of Fwd Packets

Fwd Packet Length Max

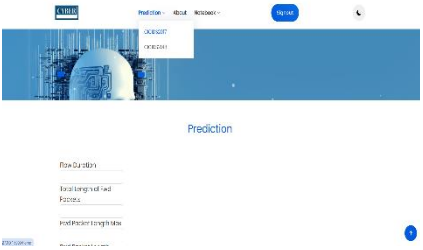
Fig.2. Upload CICIDS 2017 input values



Prediction

Result: There is an No Attack Detected, it is Normal!

Fig.3. Prediction Result



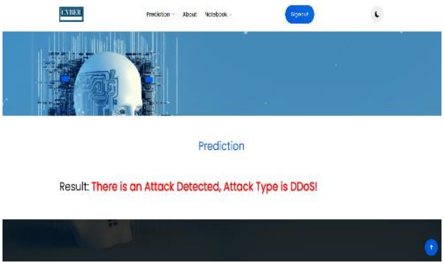
Prediction

Flow Duration

Total Length of Fwd Packets

Fwd Packet Length Max

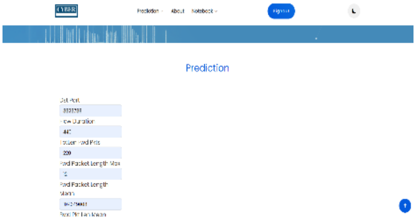
Fig.4. Upload CICIDS 2017 input values



Prediction

Result: There is an Attack Detected, Attack Type is DDoS!

Fig.5. Prediction Result



Prediction

Dst Port

[illegible]

Flow Duration

440

TotLen Fwd Pkts

220

Fwd Packet Length Max

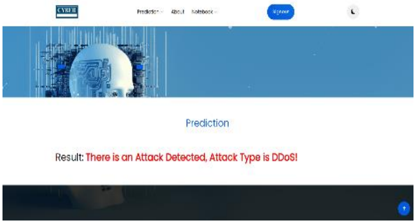
[illegible]

Fwd Packet Length Mean

[illegible]

Fwd Pkt Len Mean

Fig.6. Upload CICIDS 2018 input values



Prediction

Result: There is an Attack Detected, Attack Type is DDoS!

Fig.7. Prediction Result

So far as we can tell, the subject at hand is a research article detailing a model for the prevention and training of DDoS assaults. In order to decrease the data size, the proposed approach integrates principal component analysis with machine learning techniques. According to the authors, the four most crucial the assessment metrics. As for other criteria, they used F1-Score, accuracy, precision, and memory. The model was developed utilizing the training time. For the purpose of testing their model. The above mentioned two datasets were utilized in the study.

6 Conclusion

Improving a detection model is devised to identify DDoS attacks utilizing the IoT. sensors that can identify when a company is under heavy stress Our study primarily focuses on this aspect. We managed the AI's strength by combining several Head Part Analysis (PCA) computations, such as choice trees, irregular woodland, credulous bayes. To gauge the efficacy of our approach, we ran tests on the Two commonly used datasets are the primary focus of our study. Our review yielded excellent findings. While demonstrating great performance in preparation for DDoS assaults, our recommended strategy expedited the process of model development. Specifically, when contrasted with models that did not include PCA mix, those that made use of DT and ET calculations had very high accuracy limitations. The time needed to prepare for DT models was halved, whereas for ET models it was doubled. We want to feed our model additional real-world data in the future so that it can have more perspectives. Because of this, it will be able to adapt to different Internet of Things scenarios. To further enhance our model's compatibility with common sense Internet of Things applications, we are also diligently enhancing its multiclass grouping capabilities. So that they may be more extensively utilized in our global society, our continuing efforts demonstrate that we are still dedicated to protecting IoT devices from digital threats.

7 References

[1]. D. Velasquez, E. Perez, X. Oregui, A. Artetxe, J. Manteca, J. E. Mansilla, M. Toro, M. Maiza, and B. Sierra, ‘‘A hybrid machine-learning ensemble for anomaly detection in real-time industry 4.0 systems,’’ IEEE Access, vol. 10, pp. 72024–72036, 2022.

[2]. S. U. Rehman and V. Gruhn, ‘‘An approach to secure smart homes in cyber-physical systems/Internet-of-Things,’’ in Proc. 5th Int. Conf. Softw. Defined Syst. (SDS), Barcelona, Spain, Apr. 2018, pp. 126–129.

[3]. S. K. Vishwakarma, P. Upadhyaya, B. Kumari, and A. K. Mishra, ‘‘Smart energy efficient home automation system using IoT,’’ in Proc. 4th Int. Conf. Internet Things, Smart Innov. Usages (IoT-SIU), Ghaziabad, India, Apr. 2019, pp. 417–420.

[4]. S. Chaudhary, R. Johari, R. Bhatia, K. Gupta, and A. Bhatnagar, ‘‘CRAIoT: Concept, review and application(s) of IoT,’’ in Proc. 4th Int. Conf. Internet Things, Smart Innov. Usages (IoT-SIU), Ghaziabad, India, Apr. 2019, pp. 402–405.

[5]. (2022). Lionel Sujay Vailshery. [Online]. Available: https://www.statista. com

[6]. N. Mishra and S. Pandya, ‘‘Internet of Things applications, security challenges, attacks, intrusion detection, and future visions: A systematic review,’’ IEEE Access, vol. 9, pp. 59353–59377, 2021.

[7]. X-Force Threat Intelligence Index 2022, IBM Security, Atlanta, GA, USA, 2022.

[8]. D. Patel, ‘‘A study on DDOS attacks, danger and its prevention,’’ Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 12, pp. 1962–1967, Dec. 2022.

[9]. N. Vlajic and D. Zhou, ‘‘IoT as a land of opportunity for DDoS hackers,’’ Computer, vol. 51, no. 7, pp. 26–34, Jul. 2018.

[10].T. U. Sheikh, H. Rahman, H. S. Al-Qahtani, T. K. Hazra, and N. U. Sheikh, ‘‘Countermeasure of attack vectors using signature-based IDS in IoT

environments,'' in Proc. IEEE 10th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON), Vancouver, BC, Canada, Oct. 2019, pp. 1130–1136.

[11]. R. Zhang, J.-P. Condomines, N. Larrieu, and R. Chemali, ''Design of a novel network intrusion detection system for drone communications,'' in Proc. IEEE/AIAA 37th Digit. Avionics Syst. Conf. (DASC), London, U.K., Sep. 2018, pp. 241–250.

[12]. F. Suthar, N. Patel, and S. V. O. Khanna, ''A signature-based botnet (Emotet) detection mechanism,'' Int. J. Eng. Trends Technol., vol. 70, no. 5, pp. 185–193, May 2022.

[13]. A. M. da Silva Cardoso, R. F. Lopes, A. S. Teles, and F. B. V. Magalhaes, ''Poster abstract: Real-time DDoS detection based on complex event processing for IoT,'' in Proc. IEEE/ACM 3rd Int. Conf. Internet-Things Design Implement. (IoTDI), Orlando, FL, USA, Apr. 2018, pp. 273–274.

[14]. M. Dimolianis, A. Pavlidis, and V. Maglaris, ''Signature-based traffic classification and mitigation for DDoS attacks using programmable network data planes,'' IEEE Access, vol. 9, pp. 113061–113076, 2021.

[15]. A. Praseed and P. S. Thilagam, ''HTTP request pattern based signatures for early application layer DDoS detection: A firewall agnostic approach,'' J. Inf. Secur. Appl., vol. 65, Mar. 2022, Art. no. 103090.

[16]. X. You, Y. Feng, and K. Sakurai, ''Packet in message based DDoS attack detection in SDN network using OpenFlow,'' in Proc. 5th Int. Symp. Comput. Netw. (CANDAR), Nov. 2017, pp. 522–528.

[17]. K. Wehbi, L. Hong, T. Al-salah, and A. A. Bhutta, ''A survey on machine learning based detection on DDoS attacks for IoT systems,'' in Proc. SoutheastCon, Huntsville, AL, USA, Apr. 2019, pp. 1–6.

[18]. 2. S. Rao Polamuri, L. Nalla, A. D. Madhuri, S. Kalagara, B. Subrahmanyam and P. B. L. Aparna, "Analyse The Energy Consumption by Integrating the IOT and Pattern Recognition Technique," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024, pp. 607-610, doi: 10.1109/ICDT61202.2024.10489265

[19]. Avanija, J., K. E. Kumar, Ch Usha Kumari, G. Naga Jyothi, K. Srujan Raju, and K. Reddy Madhavi. "Enhancing Network Forensic and Deep Learning Mechanism for Internet of Things Networks." (2023).

[20]. D. Erhan and E. Anarim, ''Hybrid DDoS detection framework using matching pursuit algorithm,'' IEEE Access, vol. 8, pp. 118912–118923, 2020.