Mitigating DoS attacks in IoT using Supervised and Unsupervised Algorithms – A Survey

S.B.Gopal\textsuperscript{1*}, C.Poongodi\textsuperscript{2}, D.Nanthiya\textsuperscript{3}, R.Snega priya\textsuperscript{4}, G.Saran\textsuperscript{5}, M.Sathya Priya\textsuperscript{6}

\textsuperscript{1}Department of Electronics and Communication Engineering, Kongu Engineering college, Perundurai
\textsuperscript{2}Department of Information Technology, Kongu Engineering college, Perundurai
\textsuperscript{3}Department of CT-UG, Kongu Engineering college, Perundurai
\textsuperscript{4,5,6}Department of Electronics and Communication Engineering, Kongu Engineering college, Perundurai

*Email: gopal.ece@kongu.ac.in

Abstract—IoT is an evolving technology used in enormous applications in order to reduce the human intervention. As IoT is used in hybrid environment where it has to enable communication between multidisciplinary components it is facing lot more challenges. Security is the main issue had to be addressed. IoT devices are with limited power consumption; hence it’s not possible implement existing security algorithms as it is. Security attacks are increasing day by day; hence the dynamic solution has to be given rather than static solution. Machine learning is a promised technology that can be used in order to solve the security issues. This paper list out various work so far carried out in the area of machine learning for IoT security and solutions provided with future direction of research in this area.

Keywords—IoT, Security, Machine Learning, Network layer, DDoS attack, supervised algorithm, unsupervised algorithm

1. Introduction

IoT is an extended version of wireless sensor networks, which come into existence after the invention of Low Power Personal area network over IPV6 (PAN). The Internet of Things is used to transfer the data received from various devices deployed in remote area to its master node through Internet. IOT extends network connectivity between computers to any physical devices in our day to day life. Application area for IOT includes Smart home, smart city, smart grid, industrial internet, connected cars, smart agriculture, Environmental Monitoring and Health care. Number of devices deployed in IOT increased year over year as its application area keeps expanding. Main research issues in IOT include low power and memory, Security, Intelligent. As Memory available in sensor nodes are less only limited processing of data can be done. IoT devices are battery powered with limited energy. Because of its limited power consumption up gradation has to be done in security algorithms.

Machine learning is a promising alternative for existing security algorithms. ML is an artificial intelligence technique which performs best in dynamic networks as it is not preprogrammed. ML methods are used for training the device with various datasets to identify different attacks and to provide corresponding solution. By this way, the attacks can be detected earlier leads to protect data. By training devices with different data sets and also with live data, database will be updated with new attacks. Hence, ML algorithms can provide better security for IoT devices compared to existing static algorithms.
2. IoT Security Attacks

In IoT Security attacks will happen at three different layers namely Perception layer, Network layer and Application layer. This survey paper discuss about only about attacks related to Network layer and solution related to that in machine learning. The attacks those are possible to happen in network layer namely Denial of Service, Black Hole, Wormhole [2], Selective Forwarding, Sybil Attacks, Eavesdropping, Spoofing, Traffic Analysis, Jamming, Man in the middle, Routing attacks etc.

Cyber Attacks

Cyberattacks are the vulnerable which uses one or more computers or a program as virus, that acts against a particular device over a network. Those attacks can steal or destroy or alter the information transferred from the transmitter to the receiver by interrupting with other devices, sometimes it may disable the devices. These attacks have been mainly classified as Active attacks and Passive attacks.

2.1 Active Attacks

In active attacks [5], the system resources can be altered or damaged or stolen or deleted because the attacker accesses the system network anonymously. It affects the normal operation of the system. In this type of attack, hackers gather information from the target during passive attack. In the network layer, the major active attacks include Denial Of Service (DOS), Spoofing attacks, Sybil attacks, Jamming, Man in the Middle (MITM), Routing attack, Selective Forwarding attack, and Hole attacks.

Denial of Service (DOS)

Denial of service [3] attacks creates many unwanted requests to the system. Therefore, the actual user of the device cannot access it, as the hackers keep the server always busy. Due to this attack, the device always gets turned ON, hence the battery lifetime of IOT device gets affected. The Distributed Denial Of Service (DDOS)[29][30] can be defined as the number of systems target the single user to create more traffic.

![Figure 1 Classification of IoT Attacks](image-url)
Man-in-the-Middle (MITM)

Man in the middle attack is a system in a network where the attackers are directly coupled to the other users, creates a fake identity and fake data in the communicating system, thus exploits all the secured and original data. It manipulates and determines what data would be received by the receiver. It is a type of eavesdropping attack.

Spoofing Attacks

In the spoofing attack, the computer in a network is hacked by identifying the RFID, MAC address of the system or IP address, where the attacker pretends to be someone else to hack. It mainly includes the stealing or destroying the secured data. IP spoofing attack is the most common attack[31].

Sybil Attacks

In Sybil attack, a person creates a large number of fake identities and hence influences the systems to gain the information of the user. Those hackers steal the user privacy and spread spam. This attack also leads to DOS attacks and Man-in-the-Middle attacks.

Jamming Attacks

Jamming attacks are the subset of the DOS attacks. The unwanted signals are sent in the route of actual signal transmitted which keeps the network busy and hence the user cannot be able to access the system. It consumes more energy, more memory and degrades the device performance.

Routing Attacks

Routing attacks are the attacks by the nodes present between the receiver and the sender in a particular network. It happens when the data is being forwarded, and the hackers manipulates the whole data path. After manipulating the data path, the data can also be able to change. Routing attack also includes selective forwarding attack and hole attacks.

Selective Forwarding Attacks

Selective forwarding attack is the attacker in which the certain packets are dropped by a node during the transmission. The attacker creates a hole in the data transmitted over a communication system. These attacks are not easy to identify. These attacks aim at eroding the routing paths.

Hole Attacks

Hole attacks [6] are mentioned as follows: A Black Hole Attack is a type of DOS attack in which the total (either incoming or outgoing) traffic of the network is silently discarded without the user’s knowledge. It can be detected only when the lost traffic is monitored. A Sink Hole Attack is the attack where the hacker creates a fake information and route requests to the neighbor in a network. It results in network traffic and collapse of the network communication. When any attack combines with sinkhole attack, it is the deadliest combo. A Wormhole Attack is called as grave attack in which two attackers present in the network, records the information silently.

2.2 PASSIVE ATTACKS

Passive attacks are the attacks that take away the user’s information without the knowledge of the user and decrypt the privately secured data. Since it is unknown to the user, these types of attacks are difficult to detect. The passive attacks in the network layer include eavesdropping, traffic analysis and so on.
Eavesdropping

Eavesdropping literally means that listening to one’s word by hiding. In the networks, this attack is also known as sniffing or snooping attack, which is a theft of information that occurs in a particular network which is unsecured. In this attack, the hacker is able to capture the packets in the unsecured network before it is been transmitted to the receiver. It is of two types which are Active eavesdropping where the hacker reveals the fake identity, whereas Passive eavesdropping where the actually present is unknown.

Traffic Analysis Attack

Traffic analysis attack is defined as the hacker gets the information about the communication path in a network where the actual data gets transmitted, in which they can find the amount of data that travels throughout the path. It occurs even in the secured encrypted data, but it does not modify the data just knows the data traffic. The data are not suitable for decrypting. It is also of two types. Firstly, Active traffic analysis method makes the delay in the packet transfer time in the network. Lately, in Passive traffic analysis method, the hackers wait at both the ends of transmitter and receiver for the analysis of the traffic over a network.

3. Machine learning

Machine learning[7] is a part of artificial intelligence (AI) where the machine is not programmed explicitly but it learns automatically from the dataset given as input and hence results in the greater accuracy of the result. It uses the statistical techniques rather than programming the system[27]. The major aim of machine learning is to insist the machine to learn by itself. It actually reduces the human interaction with the machine or a system. The learning methods of machine learning can be broadly classified as follows:

- Supervised Learning
- Unsupervised Learning
- Semi-Supervised Learning
- Reinforcement Learning

Due to the increase of IOT devices in our day-to-day life, there will be numerous issues in securing the IOT devices. Machine learning will also be the best solution for the security of IOT. ML algorithms [4][28] collects the information in case of any defects in the sensors connected to the IOT devices and detects for the rectification of defects. In particular, Supervised learning can be applied to the IOT secure process.

| S. No | Author | Infrastructure | Algorithm | Dataset | Evaluation Parameters | Outcomes | Future Scope/Limitations | Year |
|-------|--------|----------------|-----------|---------|----------------------|----------|-------------------------|------|
| 1     | [8] Gini | KNN LSV | MDT | RF | NN | dataset of 491,855 packets (459,565 malicious +32,29) Packet Size, Inter-pack Interval, Bandwidth | packet-level machine learning DoS detection can accurately distinguish normal and DoS attack traffic from consumer IoT | research into machine learning anomaly detection to protect networks from insecure IoT devices limited feature set to restrict computational overhead, important for real-time classification and middlebox | 2018 |
|   |   |   |   |
|---|---|---|---|
| 2 | [9] | Proposed Model | ANN-MLP ISOT dataset, NSL-KDD dataset, feedback dataset |
|   |   |   | Accuracy proposed a dynamical MLP-based detection method against the DDoS attack through combing with sequential feature selection and feedback mechanism |
|   |   |   | deployment aim to investigate a more effective and lightweight solution to perceive the detection errors, implement our solution 2020 |
| 3 | [10] | Keras with TensorFlow backen d for deep learning and thundersvm for SVM SVM DFF DARPA 2009 dataset |
|   |   |   | Accuracy, Recall, Precision and F score SVM and DFF have been evaluated to demonstrate the feasibility of applying these algorithms. DFF can classify the data with a higher accuracy. SVM is an appropriate choice for faster classification method. to examine both two algorithms with the real time data to develop a useful method that is available with the real networks 2018 |
| 4 | [11] | - | DT, NB, ANN Raw Network Traffic Data |
|   |   |   | True Positive (TP), False Positive (FP), Recall, F-measure, Precision, RUC/AOC artificial neural network has the best accuracy rate, compared with two other methods. The accuracy rate of the artificial neural network algorithm reached 84.3% other external factors and maybe something amiss that may have contributed, so more in-depth research and testing is needed to get better results 2019 |
| 5 | [12] | - | NB, DT, LDA Kddcup9 [48], NSL-KDD [49], 2.6% of Kddcup99 |
|   |   |   | Computation times and estimated precisions of K-Fold and F-Hold results of the comparison study of 3 supervised machine learning classifiers, NB, DT, LDA suggests that the DT provides better defence against DDOS; big train datasets can fit more than small datasets analysis of F-Hold Cross-Validation will be done 2017 |
| 6 | [13] | Tensor Flow BNN, LSTM RNN datasets of CAIDA and DARPA |
|   |   |   | TPR, PPV, F1 score, Accuracy preprocessing methods, optimizers, and network architectures that are appropriate Advanced technologies of ML that pertains to time series data and newly developed statistical models for time series need to be compared 2019 |
| Page | Year | Methods | Dataset | Metrics | Challenges | Conclusion |
|------|------|---------|---------|---------|------------|------------|
| 7    | 2019 | SVM, ANN, NB, DT, USML | KDD99, UNBS-NB-15, The CAIDA UCSD Dataset | Accuracy, False Alarm Rate (FAR), Sensitivity, Specificity, False Positive Rate (FPR), AUC, Matthews correlation coefficient (MCC) | Analyzed algorithms are SVM, ANN, NB, DT, and USML. It has been shown that USML (unsupervised learning) is the best | Analysis of DDoS attacks based on the vulnerabilities of services such as Heartbleed and web brute force attack, enhancement in the multiple-class classification, self-configuration of the system, developing methods for correlating triggered alarms, and formulating protective measures. |
| 8    | 2019 | MLP, RF, KNN, K-means | ISOT dataset | Recall, Accuracy, f-measure, and precision | Both supervised learning and unsupervised learning have some advantages and disadvantages which cannot select one specific algorithm that fully cover the detection of botnets. Several other data sets may be taken into account to verify the credibility of the machine learning methods. | |
| 9    | 2018 | Mean Shift, RF, KNN, K-means | KDD99 dataset | Detection rate and accuracy | Implement a MeanShift algorithm to detect an attack in a network traffic dataset. | How to optimize that algorithm, how to examine the algorithm in larger scale of datasets, and many other gaps that still has space to work on it. |
| 10   | 2020 | Mean Shift | KDD99 dataset | Detection rate and accuracy | A hybrid model using a K-Means Algorithm and a MeanShift algorithm will be developed and tested on the KDD 99 dataset to explore improving detection accuracy. | |
| ID | JN | 
|----|----|  
| 11 | [18] | - KNN NB SVM RF ANN | Different datasets (more than 10) | accuracy, true/false positives, true/false negatives, and error | KNN algorithm exhibits the best performance overall followed by SVM algorithm, whereas low-dimensional data is better analyzed by the RF algorithm to include diversity of the machine learning algorithms, e.g., supervised, unsupervised, and semi-supervised models across multiple DDoS-related datasets | 2019 |  
| 12 | [19] | - Grey wolf optimization algorithm N-BaIoT dataset OCS VM | true positive rate, false positive rate | to detect IoT botnet attacks by utilizing GWO to optimize the hyperparameters of OCSVM and at the same time to perform feature selection to investigate the efficiency of the proposed algorithm for other types of IoT devices based on Big Data training samples | 2019 |  
| 13 | [20] | - Intrusion Detection Systems EM-CURE Cluster CAIDA DDoS Attack 2007 Dataset, The CAIDA Anonymized Internet Traces 2008 Dataset and DARPA 2000 | Accuracy, TN, FP, FN and TP unsupervised clustering algorithm (CURE), introduce a better approach to detect DDoS attacks with higher accuracy and less false alarm rate designing a proactive approach to detect DDoS attacks on cloud computing could provide an execution environment to detect and mitigate vectors of DDoS attacks | 2017 |  
| 14 | [21] | - KNIME NB DT KNN MLP RF NSL-KDD dataset | Precision Recall F-measure Accuracy Random forest in both datasets analyzing reappearance better results but in a special situation any of other algorithms may work better extend our framework to detect attack by considering unsupervised learning methods. | 2019 |  
| 15 | [22] | - Extra Trees ensemble classi NSL-KDD the UNB ISCX IDS 2012 and | FPR F-measure Cluster Purity semi-supervised DDoS detection approach based on entropy estimation, co-clustering, to evaluate its performances in real world dataset scenarios | 2018 |
The various works done so far related to Mitigating DDoS attack using supervised learning and unsupervised learning algorithm are summarized in above listed table.

4. conclusion and future work

The Security attacks are major drawback, has to be addressed prior to other problems in IoT. The Limitations and future scope of various work done are summarized in the above table. Practical Real time dataset[25] has to be considered for getting more accuracy in training and also in testing phase of machine learning algorithm. Complexity of the machine learning algorithm has to be reduced inorder to reduce the delay in detecting of attacks in real time scenario. This survey introduced different security attacks at various layers of an IoT application. Different open issues and issues that begin from the arrangement itself have likewise been examined. The best in class of IoT security has likewise been talked about with a portion of things to come research headings to upgrade the security levels is IoT. This study is required to fill in as a significant asset for security upgrade for Implementing IoT applications.

REFERENCES

[1]. Gopal S.B, Poongodi C, Joseph Auxilius Jude M (2020), Minimum energy consumption objective function for RPL in internet of things', International Journal of Scientific and Technology Research, Vol 9.
[2]. Nanthiya D, Keerthika P, Gopal S.B. (2020), Anticipation of wormhole attacks by selective routing in wireless sensor networks, International Journal of Scientific and Technology Research, Vol 9.
[3]. Thapngam T, Yu S, Zhou W, Makki S (2012) Distributed Denial of service (DDoS) detection by traffic pattern analysis. In: Peerto-Peer networking and applications December 2014, Springer,Vol 7, Issue 4, pp 346–358
[4]. Kirubavathi G, Anitha R (2018) Structural analysis and detection of android Botnets using machine learning techniques. Int J InfSecur 17(2):153–167
[5]. Tahsien S. M, Karimipour H,& Spachos P. (2020). Machine learning based solutions for security of Internet of Things (IoT): A survey. Journal of Network and Computer Applications, 102630. doi:10.1016/j.jnca.2020.102630
[6]. Butun I, Osterberz P & Sone H. (2019). Security of the Internet of Things: Vulnerabilities, Attacks and Countermeasures. IEEE Communications Surveys & Tutorials, 1–1, doi:10.1109/comst.2019.2953364.
[7]. Amit Sagu, Nasib Singh Gill(2020), “Machine Learning Techniques for Securing IoT Environment”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9.
[8]. Doshi, Rohan & Athorpe, Noah & Freamster, Nick. (2018). Machine Learning DDoS Detection for Consumer Internet of Things Devices.
[9]. Meng Wang, Yiqin Lu, Jiancheng Qin, (2020) A dynamic MLP-based DDoS attack detection method using feature selection and feedback, Computers & Security, Volume 88, 101645, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2019.101645
[10]. P. Khuphiran P. Leelaprute P. Uthayopas K. Ichikawa and W. Watanakesuontorn “Performance Comparison of Machine Learning Models for DDoS Attacks Detection,” 2018 22nd International Computer Science and Engineering Conference (ICSEC), Chiang Mai, Thailand, 2018, pp. 1-4, doi: 10.1109/ICSEC.2018.8712757.

[11]. Ahmad Sammorino (2019), “A study for [DDoS] attack classification method”, Journal of Physics: Conference Series IOP Publishing.

[12]. Wedad Alawad, Mohamed Zohdy, Debashish Deb Nath (2017), “A Study of DDoS Attacks Detection Using Supervised Machine Learning and a Comparative Cross-Validation”, www.jfitee.com Vol. 5.

[13]. Kim M. (2019). Supervised learning based DDoS attacks detection: Tuning hyperparameters. ETRI Journal, 41(5), 560-573. https://doi.org/10.4218/etrij.2019-0156.

[14]. Lima Filho F. S. de, Silveira F. A. F. de Medeiros Brito Junior, A., Vargas-Solar, G & Silveira L. F. (2019). Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning. Security and Communication Networks, 2019, 1–15. doi:10.1155/2019/1574749

[15]. Tuan T. A Long H. V, Son, L. HKumar, R Privadashini, I. & Son N. T. K. (2019). Performance evaluation of Botnet DDoS attack detection using machine learning; Evolutionary Intelligence. doi:10.1007/s12696-019-00310-w

[16]. Amirhossein Rezaei (2018), “Identifying Botnet on IoT and Cloud by Using Machine Learning Techniques”, Open International Journal of Informatics (OIJ) Vol.6

[17]. Avinash Kumar, William Bradley Glisson, Ryan Benton. 2020. “Network Attack Detection using an Unsupervised Machine Learning Algorithm”, Proceedings of the 53rd Hawaii International Conference on System Sciences

[18]. Muhammad Aamir, Syed Mustafa Ali Zaidi, “DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation”, 2019, International Journal of Information Security https://doi.org/10.1007/s10207-019-00434-1

[19]. Amaal Al Shorman,  Hossam Faris, Ibrahim Alajarah, “Unsupervised intelligent system based on one class support vector machine and Grey Wolf optimization for IoT botnet detection”, Journal of Ambient Intelligence and Humanized Computing (2019) https://doi.org/10.1007/s12652-019-01387-y

[20]. Bhaya W. & EbadyuManaa M. (2017). DDoS attack detection approach using an efficient cluster analysis in large data scale. 2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT). doi: 10.1109/ntict.2017.7976110

[21]. Hosseini S.& Azzii M. (2019). The Hybrid Technique for DDoS Detection with Supervised Learning Algorithms. Computer Networks. doi:10.1016/j.comnet.2019.04.027

[22]. Idhammad M. Afdel K.& Belouch M. (2018). Semi-supervised machine learning approach for DDoS detection. Applied Intelligence, doi:10.1007/s10489-018-1441-2

[23]. K. Narasimha Mallikarjunan, A. Bhuvaneshwaran, K. Sundarakantham and S. Mercy Shalinie, “Detecting DDoS Attacks Using Machine Learning Approach”, (2019), Theories, Applications and Future Directions—Volume I, Advances in Intelligent Systemsand Computing 798, https://doi.org/10.1007/978-81-313-1132-1_21

[24]. Zekri  M, Kafiali S. E, Aboutabiti N & Saadil Y. (2017). "DDoS attack detection using machine learning techniques in cloud computing environments. 2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech). doi: 10.1109/cloudtech.2017.8284731

[25]. Chen Liguo & ZhangYuedong & Zhao Qi & Geng Guo & Yan Zhiewei. (2018). Detection of DNS DDoS Attacks with Random Forest Algorithm on Spark. Procedia Computer Science. 134. 310-315. 10.1016/j.procs.2017.07.177.

[26]. K. Logeswaran, P. Suresh and A. R. Kannan, “Contemplate Study of Contemporary Techniques for HUIM,” 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW), Erode, India, 2018, pp. 282-286, doi: 10.1109/I2C2SW45816.2018.8997506.

[27]. K. R. Prasanna Kumar, M. Pranesh, T. G. Rahul Raajbe, P. Vignesh, K. Kousalya, and K. Logeswaran, ‘Factual product recommendation system by eliminating fake reviews using machine learning techniques’, Int. J. Adv. Sci. Technol., vol. 29, no. 7 Special Issue, pp. 2729–2735, Apr. 2020.

[28]. PCD Kalaivaani, R. Thangarajan ,Enhancing the Classification Accuracy in Sentiment Analysis using Joint Sentiment Topic Detection with Naive Bayes Classifier,Asian Journal of Research in Social Sciences and Humanities,6,12,105-116,2016,Asian Research Consortium

[29]. S.Malliga Dr.A.Tamilarasi, (2010). An autonomous framework for early detection of spoofed flooding attacks, International Journal of Network Security, Vol.10, No.1, PP.39–50.

[30]. S.Malliga Dr.A.Tamilarasi, (2009). “An Autonomous level defense for DoS/DDoS”, International Journal of Recent Trends in Engineering, Vol. 1, No. 1, pp. 518-522.

[31]. S.Malliga and Dr.A.Tamilarasi,(2008), “A distributed defensive architecture for DoS/DDoS attacks”, Journal of Information Privacy and Security, Vol. 4, No. 4, pp. 21-44.