International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

# A Survey : Deep Learning based Intrusion Detection Systems for IoT Frameworks

**Dr. Nanda Ashwin**
Professor, Department of Information Science and Engineering, Bangalore, India

**Article Info**
Volume 9, Issue 3
Page Number : 125-132

**Publication Issue**
May-June-2022

**Article History**
Accepted : 05 May 2022
Published : 20 May 2022

**ABSTRACT**

Security problems have become a key issue with the rapid rise of the Internet of Things (IoT) and its integration into numerous areas. Intrusion Detection Systems (IDSs) are critical in protecting IoT networks from potential cyber threats. Since they can automatically extract relevant properties from highly dimensional data, deep learning algorithms have drawn a lot of interest in the field of intrusion detection. This survey article gives an in-depth look at the most recent intrusion detection solutions for IoT frameworks that use deep learning approaches. We examine several approaches, methodologies, and techniques used in the design and implementation of deep learning-based intrusion detection systems. Furthermore, we examine the field's problems, open research concerns, and prospective future paths.

Keywords- DDoS Attack, Deep Learning, Recurrent Neural Network, LSTM

## I. INTRODUCTION

Network security has been significantly impacted by distributed denial of service (DDoS) attacks, making this a well-known study area. It is difficult to identify the patterns and traits of these attacks, in part because there are so many tools that are easily accessible that produce malicious traffic. Furthermore, even untrained attackers are able to initiate DDoS attacks due to the difficulties in determining the source of spoofed attack addresses. DDoS attacks' main goal is to prevent genuine users from using services, resulting in losses for the targeted companies and harm to their reputation. These assaults, sometimes referred to as protocol attacks, transmit malicious packets to the target in order to take advantage of flaws in protocols or programmes. An example of one of these attacks is The Ping of Death Attack. An additional DDoS format [1].

The use of compromised servers or botnets has made it possible for massive DDoS attacks to be launched quickly, taking only a few minutes to complete. The DDoS attacks launched by the Mirai botnet and its variants, which significantly disrupted internet infrastructure, first appeared in 2017 [2]. One noteworthy occurrence occurred in February 2018 and was known as the memcached DDoS attack. It was the greatest DDoS attack ever seen on GitHub [3]. Over time, DDoS attacks have gotten bigger, reaching 100 Gbps in 2010 [4]. As a result, it is increasingly crucial to build robust defences against DDoS attacks.

**Copyright: ©** the author(s), publisher and licensee Technoscience Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

Authorised users cannot access shared services or resources due to a denial of service (DoS) attack. A Distributed Denial of Service (DDoS) attack involves multiple attackers using diverse distributed computer resources to perform a coordinated DoS assault on one or more targets. Such attacks, spanning the System layer to the Application layer, typically concentrate on depleting network bandwidth and compromising system resources. Since the first DDoS attack occurred in 1999, DDoS has emerged as a critical, pervasive, and rapidly evolving threat worldwide. According to a survey conducted by Radware, DDoS currently ranks as the most significant threat (according to 50% of respondents in the survey) for organizations. Akamai observed 24 DDoS attack vectors in Q4 2015, representing a 148.85% increase in total DDoS attacks compared to Q4 2014, with a significant rise in multi-vector attacks. At this time, serious attack vectors include UDP flood, HTTP flood, SYN flood, ICMP flood, DNS flood, and numerous others, posing serious dangers to both systems and networks. Attacks on Distributed Denial of Service (DDoS) must be detected in order for DDoS defence strategies to be effective. However, due to attackers' attempts to mimic flash crowds and the similarity between attack traffic and actual traffic make automated DDoS detection difficult. An assault with insufficient traffic may even be mistaken for real traffic in the early stages. As a result, many professionals are looking into statistical artificial intelligence (AI) methods as a way to precisely identify DDoS attacks.

AI strategies for detecting attacks in ddos based on statistical features outperform traditional statistical methods. However, they do have several limitations:

1) They rely on extensive network expertise and experimentation with DDoS attacks to select appropriate statistical features.

2) They are often limited to detecting only one or a few specific DDoS attack vectors.

3) They require regular updates to their models and threshold values to adapt to changes in systems and attack vectors.

4) They may be vulnerable to slow attack rates, where the AI detection system may struggle to identify and respond to attacks in a timely manner.

In [32], In order to detect DDoS attacks in legitimate network traffic at the victim's end and overcome the aforementioned issues, the author offers a deep learning-based method called DeepDefense. We use a sizable dataset to train our deep learning models to address challenging recognition challenges, the UNB ISCX Intrusion Detection Evaluation 2012 DataSet (referred to as ISCX2012). In our experiments, we process two days' worth of network traffic from ISCX2012 to train both shallow AI models and our deep learning models. DeepDefense employs various neural network models, including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory Neural Network (LSTM), and Gated Recurrent Unit Neural Network (GRU). These methods have resulted in considerable performance gains in various domains when trained on large datasets. The deep learn based approach is particularly suitable for analyzing the extensive volume of network traffic. LSTM and GRU networks aid in capturing the context of network packets, especially the long and short-term patterns in DDoS attack sequences. Our experimental results demonstrate that our best deep learn based model reduces the error rate by 39.69% compared to shallow AI methods on a small dataset. In a large dataset, we achieve a reduction in the error rate from 7.517% to 2.103%. This highlights the model's ability to learn from historical network packets. Furthermore, DeepDefense outperforms shallow AI methods in terms of accuracy.

### *A method for addressing network security*

The issues is the incorporation of machine learning (ML) techniques within Software-Defined Networking (SDN). ML algorithms are employed to construct models from available data, either historical or explicit [6]. The goal of ML is to develop systems

that can learn from data without explicit programming, enabling the discovery of hidden patterns and insights. Supervised Learning, Unsupervised Learning, and Semi-Supervised Learning are the three subtypes of ML. Supervised Learning utilises what it has learned to categorise unknown data using labelled input data [1]. Decision Trees, Support Vector Machines (SVM), K-Nearest Neighbour (KNN), Random Forest, and others are popular supervised learning algorithms.

### *Network Security in Machine based Learning*

Machine learning algorithms are flexible tools that are used for a variety of tasks, including classification, prediction, and regression. Some of these algorithms differ from others in that they can recognise and construct models that can characterise or forecast unknown data by learning the underlying structure of the data without relying only on labelled data. These algorithms operate by calculating distances or similarities within the data. Examples include clustering algorithms like K-means and K-medoids, as well as dimensionality reduction techniques such as Principal Component Analysis (PCA) and Self-Organizing Maps (SOM) [6].

Semi-supervised learning algorithms use a small amount of labeled data along with a significant amount of unlabeled data during the training phase. Examples of semi-supervised algorithms include Semi-Supervised Support Vector Machine, Graph Transducer, and Gaussian Fields [6]. Thus, machine learning focuses on understanding the properties or features of a problem based on the knowledge gained from the training data [7].

Machine learning algorithms demonstrate a high detection rate for network anomaly issues. However, they have certain limitations, including the requirement for relevant data during the training stage, the need to determine the number of clusters, and considerations regarding resource utilization [1].

### *Review of Deep Learning Models*

Deep Learning (DL), a subset of Artificial Intelligence (AI), utilizes powerful and computationally intensive algorithms [6]. DL resembles the learning process of human neurons, organizing ideas in a hierarchical structure. DL algorithms build computational models with high-level abstractions using multiple layers of interconnected units. DL gained popularity with the introduction of Greedy Layer-wise Unsupervised training. Deep learning (DL) is a nonlinear, multi-neuron, multilayer neural network at its core [8]. In order to learn from the input data, this network works by concurrently stimulating several neurons with different weights. The input data is processed through a number of secret layers of neurons to produce the desired output. The precise task being addressed determines the kind, number, and arrangement of neurons and layers used in the network. While a sigmoid neuron is the basic unit, There are further activation options available. At each layer, the learned features undergo transformation and serve as input for the next layer. DL is classified into Supervised, Unsupervised, and Reinforcement Learning, based on the type of data they handle. Supervised learning requires fully labeled input data, with classification and regression being the main output tasks. Examples of supervised learning methods include convolutional neural networks (CNNs). Unsupervised learning employs datasets without labels, focusing on tasks such as dimensionality reduction, clustering, and density estimation. In Reinforcement Learning. No specific labels are present in the provided dataset. Instead, state transitions are optimised via a reward-based strategy, allowing the algorithm to learn and decide what is best to do at each stage to maximise the rewards [9].

The main distinction between Deep_Learning and Machine_Learning methods resides in the technique for processing features. While Deep Learn automatically learns and extracts features, ML depends on a domain expert to do so. While DL

algorithms are better suited for huge datasets, ML techniques perform well on small datasets. GPUs are frequently employed for increased processing since DL computations are inherently parallel [10]. Deep learning-based methods have demonstrated superior results machine learning-based techniques in various classification problems. Without requiring considerable domain knowledge, the ability of DL to automatically reduce and extract features from high-dimensional datasets helps to achieve improved accuracy.

## II. RELATED WORK

In [11], a statistical technique based on entropy is used to identify DNS Reflection Amplification attacks in DDoS attacks. The technique lessens the workload on the controller by monitoring network traffic with an orchestrator (a multi-threaded server) and a network monitor (sflow). The orchestrator module and network monitor periodically communicate to exchange status and update messages, enabling an assessment of the network's state and the determination of whether more packets are required to detect high-resolution attacks. The module calculates the average response size and entropy of the destination IP. If both the average response size and the entropy fall below a predetermined level, the network is considered to be under attack. Consequently, rate limiting is put into place by adding rules. By using complete packet inspection for high-resolution assaults and sampling approaches for low-resolution attacks, the strategy also tackles issues about flow truncation and flow reduction.

[12] uses a specified window size to analyse the entropy of newly arriving packets to all hosts in the network. In an ideal situation, it is anticipated that all hosts in the network would exhibit comparable entropy levels. The entropy, however, indicates a potential assault if it falls below a set threshold. Attack traffic is generated using Scapy.

Wang et al. in [13] employ entropy estimates of incoming packets to a given target IP on an OpenFlow edge switch to identify the presence of DDoS attacks. This strategy gives switches intelligence and improves their ability to distinguish between DDoS attacks and flash crowds. The D-ITG tool was used to generate the attack traffic for the study, which makes use of the CAIDA DoS assault 2007 dataset.

The need of taking into account several qualities when calculating entropy is emphasised in JESS [14]. In this study, entropy is calculated using both the destination IP address and Transport Layer information, such as port numbers.

To successfully differentiate between Flash Events and DDoS attacks, Sahoo et al. [16] employ General Entropy and Generalised Information Distance (GID) measurements depending on the target IP address.

To identify characteristics of a DDoS assault, [17] extracts stream table status data from the switch. By sending a "onpflowstats" request message to the switch, the controller begins regular communication. The switch then replies with measurement information, such as packet count and byte count for the associated flow entries. The controller gathers data for six certain feature values in order to build a model with an SVM classifier. This investigation was conducted under simulated conditions using Mininet. The HPing tool was employed to increase traffic to create TCP, UDP, and ICMP floods to imitate network traffic. The six features/measurements included in this study are source IP count within a given time unit, source port count within a given time unit, packet and bit standard deviation over a given time period, flow entry rate within a given time unit, and the proportion of paired flow entries. The study's findings show a detection accuracy of 95.24 percent and a false alarm rate (FAR) of 1.26%.

Two flow features—packet count and the length of a flow rule—are used in [18] to identify DDoS attacks. The strategy takes Type 1 attacks into account, which are characterised by a high volume of flows with a high packet count in each flow. Additionally, it deals with Type 2 assaults, which attempt to imitate real traffic by sending packets from a variety of spoof IP

addresses. The traffic is classified using a Linear SVM classifier as Normal, Type 1, or Type 2. For Type 1 attack traffic, the flow break value is set to 0, whereas for Type 2 attack traffic, the flows are eliminated from the flow table. In this study, the CAIDA Dataset is used to train the SVM classifier module.

Kokila et al. use an SVM classifier to accomplish multiclass classification in [19]. 1998 instances from the DARPA dataset for routine traffic and 2000 cases from the DARPA intrusion detection dataset for attack data make up the dataset used to train the classifier. The study obtains a detection accuracy of 95.11 percent for SVM-based DDoS categorization.

For network traffic analysis, a two-stage categorization approach is used in [20]. After the traffic has initially been categorised using a Naive Bayes Classifier, it is further classified using an SVM Classifier. This two-stage approach successfully lowers the false alarm rate while improving detection accuracy.

DDoS attacks are grouped together by Braga et al. [21] using an artificial neural network called a Self-Organizing Map (SOM). DDoS attacks can be effectively grouped and identified using this SOM-based methodology. These characteristics include averages for packet and byte counts per stream, stream length, the proportion of pair streams, the pace of growth of single streams, and ports. Over time, several functionalities are combined.

In [22], a trained SOM and KNN are employed for DDoS categorization. By considering the entropy of characteristics such protocol, source IP address, ports (source and destination), and packet size, the approach accurately classifies DDoS attacks.

In [3], a forecast of a future value range based on the Byte count/s and Packet count/s characteristics is made using the Pauta measure and the Weighted Moving Average (WMA) in a Gaussian distribution. their method for tracking flow, which is integrated within the data plane device, makes this forecast. The traffic is considered normal if the present readings are within the expected range. A fine-grained Machine Learning technique combining an Autoencoder and a softmax classifier is then applied at the control plane if the results stray from the expected range. This algorithm enables the classification of attacks in real-time using extracted traffic data.

In [2], Network Function Virtualization (NFV) and Machine Learning (ML) techniques are combined to identify DDoS attacks in systems using Software-Defined Networking (SDN). Data plane devices use virtual network functions (VNFs) to track traffic and look for network intrusions. One of the VNFs is in charge of explicitly collecting feature data from the stream and sending it to the controller. The controller then develops a botnet attack model based on the Random Forest algorithm. By utilising the capabilities of Virtual Network Functions, this strategy intends to obtain real-time network traffic information, minimising reliance on data which is historical alone.

In [23], Wang et al. provide a probabilistic representation for identifying Low Rate Distributed Denial of Service (LDDoS) assaults using the Renyi entropy of IP addresses (both source and destination) and the Hidden Markov Model (HMM). The Viterbi technique is used to decode the HMM, and Euclidean distance is used to calculate Renyi entropies.

Four alternative machine learning (ML) techniques—K-nearest neighbour (KNN), Naive Bayes, K-means, and K-medoids—are compared in terms of detection accuracy and processing durations in [24]. Naive Bayes, which has the highest detection accuracy, with a rate of 94%.

In their study [26], Li et al. developed a deep learning defensive model with the goal of successfully reducing DDoS attack traffic. The model incorporates the Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Convolutional Neural Network (CNN) architectures. The Keras Deep Learning Framework on a GPU platform is used in this study. Notably, there hasn't been much study on applying deep learning algorithms to identify DDoS attacks while Software-Defined Networking (SDN) is being used. But as demonstrated by [27–31], a number

of DL-based initiatives have been made for Network Intrusion Detection in SDN environments. Another project, DeepDefense [32], builds a model specifically designed for employing recurrent neural networks to detect DDoS assaults in traditional networks (LSTM), convolutional neural networks (CNN), and gated recurrent units (GRU). The UNB ISCX Intrusion Detection Evaluation 2012 Dataset, a sizable dataset, is used in their research to assess the effectiveness of their methodology. They also contrast the effectiveness of their approach and the Random Forest algorithm.

To identify Application Layer DDoS assaults, [33] integrates a Stacked Autoencoder deep learning architecture. Web server logs are gathered for the study, and pertinent features are extracted from the logs. A feature normalisation strategy employing the Min-Max algorithm is used to guarantee compatibility between features. The final classification of network data is then performed using Logistic Regression after a Deep Learning Model has been used to capture dynamic information.

In [34], a Restricted Boltzmann Machine (RBM) of the Gaussian-Bernoulli type is used to build a DDoS detection model. Three deep learning algorithms—Bernoulli-Bernoulli, Gaussian-Bernoulli, and Deep Belief Network—as well as three machine learning models—Decision Tree, SVM-Epsilon SVR, and Radial Basis—are all thoroughly compared in the paper. The study assesses how well these methods work for spotting DDoS attacks. A table that lists the various DDoS detection methods and the traffic attributes used for detection is also provided.

## III. CONCLUSION

DDoS assaults have serious repercussions and, if they are not effectively mitigated, can cause major interruptions in network operations. Since these attacks are constantly changing, it is difficult for conventional defence methods to successfully thwart them. Deep learning techniques are being developed to create Knowledge Defined Networks to handle these problems. With regard to a variety of categorization issues pertaining to network security, these sophisticated algorithms are proven to be more efficient than traditional machine learning techniques. Organisations can improve their defences against sophisticated DDoS assaults and lessen the effect they have on network operations by utilising deep learning.

## IV. REFERENCES

[1]. Jing, Xuyang, Zheng Yan, and Witold Pedrycz, Security Data Collection and Data Analytics in the Internet: A Survey. IEEE Communications Surveys Tutorials (2018)

[2]. Park, Younghee, Nikhil Vijayakumar Kengalahalli, and Sang -Yoon Chang, Distributed Security Network Functions against Botnet Attacks in Software-defined Networks

[3]. Famous DDoS attacks, https://www.cloudflare.com/learning/ddos/famous-ddos-attacks

[4]. Han, Biao, Xiangrui Yang, Zhigang Sun, Jinfeng Huang, and Jinshu Su. OverWatch: A Cross-Plane DDoS Attack Defense Framework with Col-laborative Intelligence in SDN. Security and Communication Networks 2018

[5]. Gkountis, Christos, et al., Lightweight algorithm for protecting SDN controller against DDoS attacks. Wireless and Mobile Networking Conference (WMNC), 2017 10th IFIP. IEEE, 2017.

[6]. Sultana, Nasrin, et al., Survey on SDN based network intrusion detection system using machine learning approaches. Peer-to-Peer Networking and Applications (2018): 1-9.

[7]. Buczak, Anna L., and Erhan Guven. , A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys Tutorials 18.2 (2016): 1153-1176.

[8]. Fadlullah, Zubair, et al. , State-of-the-art deep learning: Evolving machine intelligence toward tomorrows intelligent network traffic control systems. IEEE Communications Surveys Tutorials 19.4 (2017): 2432-2455.

[9]. Hatcher, William Grant, and Wei Yu. , A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends. IEEE Access 6 (2018): 24411-24432.

[10]. Xin, Yang, et al. , Machine Learning and Deep Learning Methods for Cybersecurity. 2018 IEEE Access.

[11]. Zaalouk, Adel, et al. , Orchsec: An orchestrator-based architecture for enhancing network-security using network monitoring and sdn control functions. 2014 IEEE Network Operations and Management Symposium (NOMS).

[12]. Mousavi, Seyed Mohammad, and Marc St-Hilaire. , Early detection of DDoS attacks against SDN controllers. International Conference on Computing, Networking, and Communications (ICNC), 2015 IEEE

[13]. Wang, Rui, Zhiping Jia, and Lei Ju. , An entropy-based distributed DDoS detection mechanism in software-defined networking. 2015 Trust-com/BigDataSE/ISPA IEEE, 2015, Volume 1

[14]. Kalkan, Kbra, et al., JESS: Joint Entropy-Based DDoS Defense Scheme in SDN.IEEE Journal on Selected Areas in Communications 36.10 (2018): 2358-2372.

[15]. Mao, Jiewen, Weijun Deng, and Fuke Shen. , DDoS Flooding Attack Detection Based on Joint-Entropy with Multiple Traffic Features. 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). IEEE, 2018.

[16]. Sahoo, Kshira Sagar, Mayank Tiwary, and Bibhudatta Sahoo. , Detection of high rate DDoS attack from flash events using information metrics in software defined networks. Communication Systems Networks (COMSNETS), 2018 10th International Conference on. IEEE,2018.

[17]. Ye, Jin, et al. , A DDoS Attack Detection Method Based on SVM in Software Defined Network. Security and Communication Networks 2018 (2018).

[18]. Phan, Trung V., et al. , OpenFlowSIA: An optimized protection scheme for software-defined networks from flooding attacks. Communications and Electronics (ICCE), 2016 IEEE Sixth International Conference on. IEEE, 2016.

[19]. Kokila, R. T., S. Thamarai Selvi, and Kannan Govindarajan. , DDoS detection and analysis in SDN-based environment using support vector machine classifier. Advanced Computing (ICoAC), 2014 Sixth International Conference on. IEEE, 2014.

[20]. Khemapatapan, Chaiyaporn. , 2-Stage Soft Defending Scheme Against DDoS Attack Over SDN Based on NB AND SVM.

[21]. Braga, Rodrigo, Edjard Mota, and Alexandre Passito. , Lightweight DDoS flooding attack detection using NOX/OpenFlow. Local Computer Networks (LCN), 2010 IEEE 35th Conference on. IEEE, 2010.

[22]. Nam, Tran Manh, et al. , Self-organizing map-based approaches in DDoS flooding detection using SDN. 2018 International Conference on Information Networking (ICOIN). IEEE, 2018.

[23]. Wang, Wentao, Xuan Ke, and Lingxia Wang , A HMM-R Approach to Detect L-DDoS Attack Adaptively on SDN Controller. Future Internet 10.9 (2018): 83

[24]. Barki, Lohit, et al. , Detection of distributed denial of service attacks in software defined networks. Advances in Computing, Communications and Informatics (ICACCI), 2016 International. Conference on. IEEE, 2016.

[25].Niyaz, Quamar, Weiqing Sun, and Ahmad Y. Javaid. , A deep learning based DDoS detection system in software-defined networking (SDN). arXiv preprint arXiv:1611.07400 (2016).

[26].Li, Chuanhuang, et al. , Detection and defense of DDoS attackbased on deep learning in OpenFlowbased SDN. International Journal of Communication Systems 31.5 (2018): e3497.

[27].Tang, Tuan A., et al. , Deep learning approach for network intrusion detection in software defined networking.Wireless Networks and Mobile Communications (WINCOM), 2016 International Conference on.IEEE, 2016.

[28].Wang, Wei, et al. , HAST-IDS: learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection IEEE Access 6 (2018): 1792-1806.

[29].Ishitaki, Taro, et al. , Application of Deep Recurrent Neural Networks for Prediction of User Behavior in Tor Networks. Advanced Information Networking and Applications Workshops (WAINA), 2017 31st International Conference on. IEEE, 2017.

[30].Javaid, Ahmad, et al. , A deep learning approach for network intrusion detection system. Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2016.

[31].Tang, Tuan A. , et al., Deep recurrent neural network for intrusion detection in sdn-based networks. 2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft). IEEE, 2018.

[32].Yuan, Xiaoyong, Chuanhuang Li, and Xiaolin Li. , DeepDefense: Identifying DDoS Attack via Deep Learning. 2017 IEEE International Conference on Smart Computing (SMARTCOMP). IEEE, 2017.

[33].Yadav, Satyajit, and Selvakumar Subramanian. , Detection of Application Layer DDoS attack by feature learning using Stacked AutoEncoder. Computational Techniques in Information and Communication Technolo- gies (ICCTICT), 2016 International Conference on. IEEE, 2016.

[34].Imamverdiyev, Yadigar, and Fargana Abdullayeva. , Deep Learning Method for Denial of Service Attack Detection Based on Restricted Boltzmann Machine. Big Data 6.2 (2018): 159-169.

**Cite this article as :**

Dr. Nanda Ashwin, "A Survey : Deep Learning based Intrusion Detection Systems for IoT Frameworks ", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 3, pp. 125-132, May-June 2022.
Journal URL : https://ijsrst.com/IJSRST2411145