Lightweight Intrusion Detection System In IoT Networks Using Raspberry pi 3b+

Kuthada Mohan Sai*[1], Brij .B Gupta[2], Ching-Hsein HSU*[3], Dragan Peraković*[4]

[1]*Department of Computer Engineering, National Institute of Technology Kurukshetra, India*
[2]*National Institute of Technology Kurukshetra, Kurukshetra, Haryana 136119, India, & Asia University, Taichung 413, Taiwan & Staffordshire University, Stoke-on-Trent ST4 2DE, UK*
[3] *Department of Computer Science and Information Engineering, Asia University, Taiwan &. Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan & Department of Medical Research, China Medical University Hospital, China Medical University, Taiwan*
[4]*University of Zagreb, Croatia*
*Corresponding Author

Abstract

The Internet of Things (IoT) has become a new paradigm by integrating a variety of applications like healthcare, transport, manufacturing and supply chain management. This resulted in the generation of various networks which are susceptible to Cyberattacks. The most prominent attack in the Cyberattacks is Denial of Service(DoS), in this attack the attackers flood the network with huge volumes of data or requests so that it restricts the nodes from accessing various kinds of services offered in that network. Intrusion detection systems (IDS) have proved as promising defence mechanisms to detect an attack in the cyber world. However, the resource constraints like lower processing power and low power consumption in IoT devices stood as a challenge in implementing the conventional IDS techniques in IoT devices. In this paper, we implemented a lightweight Intrusion detection technique using the machine learning based approach implemented on a Raspberry Pi. In which a support vector machine (SVM) classification algorithm is utilized for detecting the adversaries in the networks and the correlation-based feature selection algorithm is utilized for selecting the features to make the model lightweight. The experimental results have shown that the attacks detection rate and accuracy are satisfactory for our approach in case of a DoS attack.

Keywords

IoT, DDoS, Machine learning, SVM, Intrusion detection

1. Introduction

In today's world many of the services and applications around us are controlled by machines with minimal human intervention this is due to the integration of the internet with many applications. The paradigm that is offering a wide range of services is the Internet of Things (IoT) it has integrated many physical objects such as sensors, cameras, vehicles, humans, healthcare. The services or the applications that are offered by an IoT can either be simple on/off automations to complete control of smart grids. In many of the IoT networks, the primary sensor nodes or the processing node are resources constrained devices such as low memory and low power consumption devices these nodes perform the data collection, data transmission, and data processing [1]. The data from the physical environment is collected by the sensor nodes and it is transmitted over the communication network which is connected from any of the following technologies like Wi-Fi, ethernet, or any other wired technologies[25]. For connecting the users and objects across distances these communication technologies are combined with the internet protocol. The data is processed by the applications to extract useful information and takes decisions to perform some actions in the physical environment.

International Conference on Smart Systems and Advanced Computing (Syscom-2021), December25–26, 2021
EMAIL: mohankuthada@gmail.com, gupta.brij@gmail.com, robertchh@gmail.com, ; dperakovic@fpz.unizg.hr

As IoT services are of a wide range they are heterogeneous and the communication technologies among them are distributed with different standards, which enlarge the threats in end-to-end security[6]. The attack surface in IoT is increased because of the expansions in complexity and diversity of the IoT[5], [17],[24].

The defence mechanisms like signature-based intrusion detection methods will not function properly for the modified attacks. Hence intrusion detection systems(IDS) for novel intrusions is essential. The Anomaly based intrusion detection scheme serves the purpose as this detection scheme does not need any prior information about the attack signatures. A detailed classification of the intrusion detection and prevention systems is given in [2][8].

Specific characteristics of IoT networks that are given by many researchers in the aspect of developing a IDS are as follows [1],[3],[27].

The nodes or the devices in an IoT network are resource constrained and run on low power hence it is not feasible to host a conventional IDS which requires high power and high computational capabilities.

The protocols used in the IoT networks are not the same as the protocols used in conventional networks. The protocols used in these networks like IPv6 over Low-power Wireless Personal Area Network (6LoWPAN) and Constrained Application Protocol (CoAP) make the network heterogeneous which results in new weaknesses and makes it challenging to implement the IDS in the networks.

Taking the above characteristics into consideration it is necessary to develop lightweight IDS that can perform efficiently in securing the networks. The term lightweight signifies that the system should work on constrained resources available in the nodes of the IoT networks and it does not signify that the system should be simple. “Roesch et al. [4] defined lightweight IDS as small, powerful, and flexible enough to be used as a permanent element of the network security infrastructure”.

Keeping the constraints in mind we propose a lightweight intrusion detection system with fewer features that are available on the datasets like DARPA, KD’99, UNSW-NB15[7]. In our approach, we used UNSW-NB15 for training the lightweight model and it is evaluated on a “Raspberry Pi 3b+ which is a credit card sized computer” and runs on low power approximately 10W which can be deployed in the IoT network. Detailed information regarding the Raspberry pi is given in [9].

The Distributed Denial of Service (DDoS) attack is one of the major threats for vehicles present in the IoV[13]. Among the taxonomy of DDoS attacks flooding attacks are of major concern as they exhaust the cache and the computing resources of the OBU present on the vehicles. DDoS attack creates huge traffic using the DDoS attack from which the attacker exhausts the resources of the targeted host like the network bandwidth and the CPU time[21].

Some of the examples of DDoS attacks are DNS flood, SYN Flood, UDP flood and Ping Flood [3]. The OBU of the vehicles is connected to various devices through wired and wireless means where each and every component is connected globally.

As the devices are connected globally and the resources of these devices have highly controlled the security of IoV becomes highly challenging. In our paper, we implement an anomaly-based lightweight DDoS detection, we generated a dataset using the OMNET++[18], sumo[19] veins [22] and INET [20] tools in order to generate a dataset for the UDP flood attack and for reducing the number of features and an efficient feature selection algorithm Correlation-based feature selection algorithm is used to train a machine-learning algorithm Support vector machine(SVM) and the J48 classifier to classify the incoming traffic as positive or negative. Many researchers have proved that

the SVM classifier has outperformed other classifier algorithms like k- nearest neighbour, random forest and even neural networks [4] in order to verify this claim we used the J48 classifier to compare it with SVM.

2. Related work

The IDS proposed in this literature [3] is based on a fine state machine or simply it can be regarded as automata. The transitions performed by the automata are used for characterizing the network and these transitions are used for detecting the intrusions. The proposed approach is evaluated for only three attacks: jamming, replay and false attack.

The model proposed in this literature [12] uses the public IDS tools called snort and bro for detecting the intrusions. These tools are installed on a Raspberry pi. However these tools can only be used with limited rules and if the number of rules increases the system gets crashed. Hence this approach is not satisfactory for practical usage.

The model proposed in this literature [14] uses a SVM based classifier and the features are selected by a hybrid feature selection algorithm in which it uses genetic algorithms and mutual information and tried to improve the accuracy of classification. Authors of this approach have proved that the SVM based classifier performs better than an artificial neural network.

The authors in this literature [15] proposed an energy efficient IDS approach based on auto aggressive mode and game theory and obtained satisfactory results.

The authors of this model [15] used the Correlation based feature selection for reducing the number of features to 7 and performed the classification using a machine learning algorithm called J48 classifier on a Raspberry Pi and achieved satisfactory results.

The support vector machine based IDS proposed in [18] uses 3 features that are extracted from the packet arrival rate. Performance evaluation of this approach is done on the on a MatLab simulation and achieved satisfactory results for the DoS attack.

The model proposed in this literature [19] has a two step mechanism for detecting the intrusions in the first step; it utilizes many numbers of binary classifiers for classifying the samples. If step -1 gives an ambiguous output, the sample is again classified by the k- nearest neighbors' algorithm. The detection rate of this algorithm is 94% but it requires high energy and more computation power as more number of classifiers are used in this approach.

3. Motivation

In general, the most widely used intrusion detection systems are anomaly and signatures based detection schemes are used and more signature based approaches are used as public IDS these attacks can only detect known attacks and are ineffective on modified known attacks or new attacks [10]. The feature of anomaly is to detect network traffic deviations from normal behaviour these features aid the IDS for detecting known as well as unknown attacks. As the IDS functions on detection of an anomaly, it might raise normal traffic as an intrusion and generate a false alarm which makes it impractical for some use cases. As the IoT devices are resource constrained most of the existing schemes are signature based IDS with limited rules. In our project, we implement an anomaly based lightweight Intrusions detection system and for reducing the number of features we utilized a correlation-based feature selection(CFS) algorithm these features are used to train a machine-learning algorithm Support vector machine(SVM) to classify the intrusions as positive or negative intrusions. Many researchers have proved that the SVM classifier has outperformed other classifier algorithms like k- nearest neighbor, random forest and even neural networks [11]. The proposed model is lightweight that it can run on a Raspberry Pi 3b+.

4. Proposed approach

To implement our model we used a Raspberry pi 3b+ running on Raspbian Operating System and used an open source performance evaluator called Weka [26] to evaluate the model performance. The SVM based classifier is used for classification by using the features extracted by the CFS algorithms. The data set utilized for training and testing is UNSW-NB 15. The proposed model architecture is given in the Figure 1.

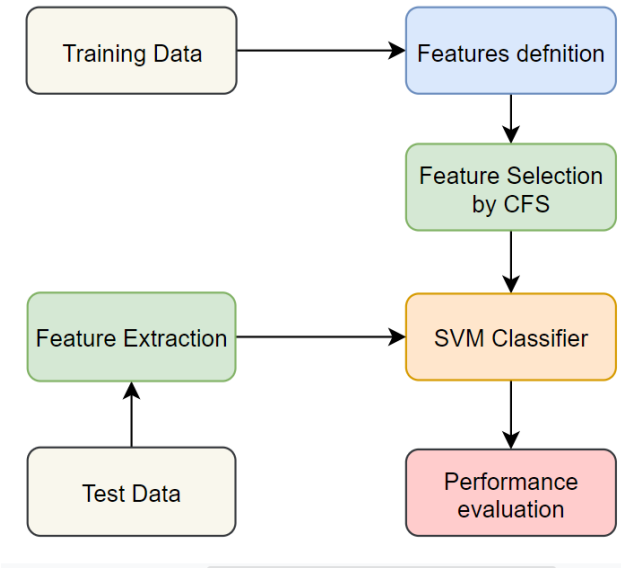

Figure 1: Proposed model Architecture

"The Dataset UNSW-NB15 was created in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) "[26]. There are 175,341 instances for nine attacks and the description of each attack is given in Table 1.

Table 1
Description of Attacks used in the UNSW-NB 15 dataset

Attack type	Description
Denial of service(DoS)	A malicious attack which is done on a host machine connected to the internet suspending its services so that the network or the server is not available to the legitimate users.
Worms	A self replicating malicious program which is used to exhaust the resources of a system. It spreads itself to other systems using the network based on the system security failures which helps the worm to access it
Shellcode	To exploit the software vulnerabilities, a small piece of code is utilized as a payload.
Reconnaissance	It can simulate all the attacks which gather information
Fuzzers	It uses randomly generated data for suspending the services offered by a

	program or by a network
Backdoors	A stealth way of accessing a system or its data by bypassing its security frameworks
Generic	With known block and size of the key, the block ciphers are deciphered by not taking care about the structure of the block.
Analysis	The technique which is used to perform the port scanning and file penetration attacks.
Exploits	The attackers know the vulnerabilities present in a Software or an OS and exploit them.

4.1. Correlation based feature selection

The selection of features is an important function in designing a machine learning model. The features that are selected should be relevant so that the designed model gives required results and they decide the performance of the model. "There are mainly three types of feature selection methods such as filter-based, wrapper-based and embedded-based approaches" [20]. These approaches have their pros and cons. The filter based approach is computationally lightweight but it is less accurate as it ignores the underlying algorithm's nature. The wrapper based approach is computationally heavy as the learning process and feature selection are combined in its process. In the third approach the training part is embedded with the feature selection.

For implementing lightweight IDS the feature selection algorithm should also be lightweight; hence we utilized the CFS algorithm which is a filter based approach. "The algorithm evaluates the relation between output and correlated inputs" [22]. The features for a classifier machine learning model can be selected based on the correlation among the features. The irrelevant and redundant features are ignored by this algorithm. The redundancy of a feature is determined by the algorithm when it is highly correlated with one or more features. The subset of features are selected when they are highly related with class and not correlated with each other. To improve the accuracy and to reduce the computation of the machine learning model the features that are redundant and not relevant are ignored. The subset of features with highest merit is selected as the feature set and these features are used for training the model. The merit of a feature subset is given by the equation 1. In the equation the number of feature present in a given subset is defined by k, the average of feature-class correlation is given by $\bar{r}_{cf}$ and average of feature-feature correlation is given by $\bar{r}_{ff}$.

$$Ms_k = \frac{k\bar{r}_{cf}}{\sqrt{k+k(k-1)\bar{r}_{ff}}} \quad \textbf{(1)}$$

4.2. Support vector machine

"The support vector machine (SVM) is a supervised machine learning algorithm which is used for both classification and regression purposes" [23]. Wide ranges of applications use SVM as a classifier. The working of SVM is based on the support vectors and hyperplane.

A hyperplane is a flat subspace that has one less dimensions of the coordinate system that it is represented. For a 2-d coordinate system it is given by equation 2

$$P_0 + P_1X_1 + P_2X_2 = 0 \quad \textbf{(2)}$$

In m-D coordinate it is represented by equation 3

$$P_0 + P_1X_1 + P_2X_2 + \ldots\ldots + P_mX_m = 0. \quad \textbf{(3)}$$

The algorithm has data points which are called support vectors and the hyperplane separating them is called SVM. The data points which are nearer to the hyper plane are called support vectors and if these points are removed from the dataset it changes the position of the hyperplane. There are many hyper planes between the support vectors but the plane with maximum is selected.

SVM has two main objectives to achieve good classification: maximize the margin and correctly classified data points. There is a tradeoff between these two to regulate this trade off the SVM has three hyper parameters they are Kernel, C (error rate) it a penalty for the data points that are classified wrongly. Gamma is the coefficient of kernels which is used for fitting the plane. If the value gamma is high the plane results in over fitting.

Not all the datasets are linear separable by the hyperplane. The SVM uses various kernels to separate non linear data point to linear separable. Kernels like rbf, poly etc project the data to higher dimensions to separate them by hyper plane and again they are projected back to the lower plane.

In our approach we used RBF (radial basis function) given in equation 4 as the kernel for the SVM and the parameter γ gamma defines the linearity of the hyper plane. If the value gamma is high the plane results in over fitting.

$$A(x_i,x_j)=\exp(-\gamma||x_i-x_j||2) \quad \textbf{(4)}$$

5. Implementation

To implement our model we used a Raspberry pi 3b+ running on Raspbian Operating System and used an open source performance evaluator called Weka [26] to evaluate the model performance. The SVM based classifier is used for classification by using the features extracted by the CFS algorithms. The data set utilized for training and testing is UNSW-NB 15.

For efficient model building it is recommended to use all the 44 features in the dataset for classifying the attacks but in our experiment we used Top 3 features selected by the CFS and are shown in table 2.

Table 2
Features selected by the CFS algorithm

Feature name	Type	Description
sbytes	Integer	The transaction bytes from source to destination.
rate	Float	The rate at which number of packets arrive
sttl	Integer	The value of time to live from source to destination.

The designed model performance is evaluated utilizing weka evaluation tool and DoS attack test dataset is used for evaluating its performance for DoS attack. When the CFS algorithm was not used the raspberry pi was not able to handle the data set and the system got crashed. So we evaluated the performance of non CFS model on a laptop and used a DoS test dataset which had 4045 instances of DoS attack and 12328 instances of normal traffic. After using the CFS algorithm the features are reduced to 3 and all the instances of the training data set were handled by the Raspberry Pi and the same test dataset is used for evaluating the model on the weka tool by using the SVM classifier. The kernel used in the SVM classifier is rbf kernel with C= 10 and gamma = 0.001 a grid search was performed on the model to find out the best training parameters for our model. The performance parameters of the model with and without using the CFS are compared and observed. The parameters used for evaluating are the confusion matrix, F-measure, recall, precision, False positive rate (FPR), True positive rate (TPR), False Negative rate (TNR), True Negative rate (TNR) and accuracy.

6. Results and observations

6.1. Results

The Confusion matrix of the proposed model by using CFS is given in Figure 2 and by not using CFS is given in Figure 3, where positive is given for the attack instance. From the confusion matrix we can observe that the model which is built using the features selected by the CFS performs on par with the model build using all the features.

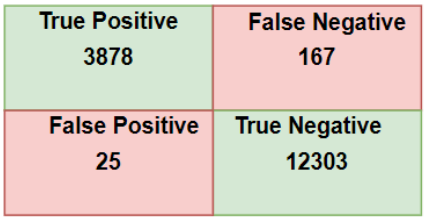

Figure 2: Confusion matrix using CFS

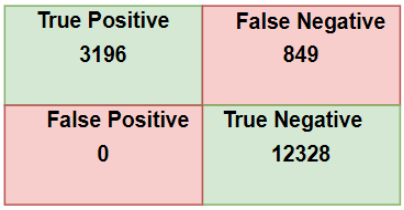

Figure 3: Confusion matrix without using CFS

The detection accuracy of the model using CFS is compared with model that has not used CFS are compared in the T4able 3.

Table 3
Detection accuracy of proposed model with and without using CFS

Parameter	With CFS	Without CFS
TPR	0.96	0.79
FPR	0.02	0
TNR	0.99	1
FNR	0.04	0.21
Precision	0.99	1
Recall	0.96	0.79
F measure	0.97	0.88
Accuracy	0.98	0.94

6.2. Observations

- By using CFS we have achieved the lightweight IDS by greatly reducing the number of feature. It is known that if the number of features are more the system becomes more complex and it

requires more computational power. By reducing the number of features from 44 to 3 we can observe that the complexity is greatly reduced and hence our system uses less power for computation.

- In the non CFS model the classifier got confused because of using more number of features and gave less number of true positives when compared to the CFS model.
- The usage of CFS algorithm reduced the complexity of the model
- The detection accuracy of the model which has used CFS is almost similar to the model which has used all the features. Hence our system achieved lightweight detection without compromising on the detection accuracy.

7. Conclusion and future plan

It is known that due to the wide range of usage of IoT devices the attacks performed on these devices increasing exponentially and these devices or nodes are highly resource constrained they cannot host defence mechanisms that were developed for the conventional systems that utilize high computation resources and power hence it is necessary to develop lightweight intrusion detection systems to safeguard the IoT devices. The signature based IDS are widely proposed for the IoT network but nowadays researchers are aiding the machine learning techniques for designing anomaly approaches. In our paper, we used a Support vector machine for classifying the attack traffic and the normal traffic. In order to achieve this in a lightweight scenario, a correlation based feature selection algorithm is used to reduce the number of features a dataset and the dataset used was UNSW-NB 15 and the features are reduced from 44 to 3, these three features are used to train the model on the entire dataset. In order to compare the CFS model with the non-CFS model, the non-CFS model is evaluated on a laptop as it is impossible to train the model using all the instances with 44 features on a raspberry pi which resulted in a system crash. The DoS attack instances are extracted from the training set of the UNSW-NB 15 and performed the evaluation using the WEKA evaluation tool. The models were compared according to their detection accuracy and it is evident that CFS algorithm has made the system lightweight by reducing the number of features. In our approach the model is only evaluated for the DoS attack and in future work, we will be evaluating a wide range of attacks and with various kinds of feature selection algorithms and with and with various kernels of the SVM for better detection accuracy.

References

[1] "B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, ''A survey of intrusion detection in Internet of Things,'' J. Netw. Comput. Appl., vol. 84, pp. 25–37, Apr. 2017.

[2] Bijone, M. (2016). A survey on secure network: intrusion detection & prevention approaches. *American Journal of Information Systems*, *4*(3), 69-88.

[3] Y. Fu, Z. Yan, J. Cao, and O. Koné, and X. Cao, ''An automata based intrusion detection method for internet of things,'' Mobile Inf. Syst., vol. 2017, May 2017, Art. no. 1750637

[4] M. Roesch et al., ''Snort—Lightweight Intrusion Detection for Networks,'' in Proc. Lisa, vol. 99, 1999, pp. 229–238.

[5] Gou, Z., Yamaguchi, S., & Gupta, B. B. (2017). Analysis of various security issues and challenges in cloud computing environment: a survey. In Identity Theft: Breakthroughs in Research and Practice (pp. 221-247). IGI global.

[6] AlZu'bi, S., Hawashin, B., Mujahed, M., Jararweh, Y., & Gupta, B. B. (2019). An efficient employment of internet of multimedia things in smart and future agriculture.

[7] I. Sharafaldin, A. Gharib, A. H. Lashkari, and A. A. Ghorbani, ''Towards a reliable intrusion detection benchmark dataset,'' Softw. Netw., vol. 2017, no. 1, pp. 177–200, 2018.

[8] S. Rizou and K. E. Psannis (2021), Enhanced Privacy Recommendations According to GDPR in the Context of Internet-of-Things (IoT), Insights2Techinfo, pp.1

[9] Raspberry Pi 3b+ datasheet. Retrieved May 15, 2021 from https://static.raspberrypi.org/files/product-briefs/Raspberry-Pi-Model-Bplus-Product-Brief.pdf
[10] Kumar, R., & Sharma, D. (2018, July). HyINT: signature-anomaly intrusion detection system. In *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.
[11] S. U. Jan, V.-H. Vu, and I. Koo, ''Throughput maximization using an SVM for multi-class hypothesis-based spectrum sensing in cognitive radio,'' Appl. Sci., vol. 8, no. 3, p. 421, 2018
[12] Dhanabal, L., & Shantharajah, S. P. (2015). A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, *4*(6), 446-452
[13] Dhananjay Singh (2021) Captcha Improvement: Security from DDoS Attack, Insights2Techinfo, pp.1
[14] H. Sedjelmaci, S. M. Senouci, and T. Taleb, ''An accurate security game for low-resource IoT devices,'' IEEE Trans. Veh. Technol., vol. 66, no. 10, pp. 9381–9393, Oct. 2017.
[15] Feng, Y., Hori, Y., Sakurai, K., & Takeuchi, J. I. (2013). A behavior-based method for detecting distributed scan attacks in darknets. *Journal of information processing*, *21*(3), 527-538
[16] Soe, Y. N., Feng, Y., Santosa, P. I., Hartanto, R., & Sakurai, K. (2019, March). Implementing Lightweight IoT-IDS on Raspberry Pi Using Correlation-Based Feature Selection and Its Performance Evaluation. In *International Conference on Advanced Information Networking and Applications* (pp. 458-469). Springer, Cham
[17] Ab Malek, M. S. B., Ahmadon, M. A. B., Yamaguchi, S., & Gupta, B. B. (2016, October). On privacy verification in the IoT service based on PN 2. In 2016 IEEE 5th Global Conference on Consumer Electronics (pp. 1-4). IEEE.
[18] Jan, S. U., Ahmed, S., Shakhov, V., & Koo, I. (2019). Toward a lightweight intrusion detection system for the internet of things. *IEEE Access*, *7*, 42450-42471.
[19] L. Li, Y. Yu, S. Bai, Y. Hou, and X. Chen, ''An effective two-step intrusion detection approach based on binary classification and κ-NN,'' IEEE Access, vol. 6, pp. 12060–12073, 2018.
[20] Hira, Z. M., & Gillies, D. F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in bioinformatics*, *2015*.
[21] Gupta, B. B., & Quamara, M. (2021). A taxonomy of various attacks on smart card–based applications and countermeasures. Concurrency and Computation: Practice and Experience, 33(7), 1-1.
[22] Hall, M. A. (1999). Correlation-based feature selection for machine learning.
[23] Rossi, F., & Villa, N. (2006). Support vector machine for functional data classification. *Neurocomputing*, *69*(7-9), 730-742.
[24] K. Yadav (2021) Blockchain for IoT Security, Insight2Techinfo, pp.1
[25] Jain, A. K., & Gupta, B. B. (2019). Feature based approach for detection of smishing messages in the mobile environment. Journal of Information Technology Research (JITR), 12(2), 17-35.
[26] Moustafa, N., & Slay, J. (2015, November). The significant features of the UNSW-NB15 and the KDD99 data sets for network intrusion detection systems. In *2015 4th international workshop on building analysis datasets and gathering experience returns for security (BADGERS)* (pp. 25-31). IEEE.
[27] Gupta, S., & Gupta, B. B. (2018). A robust server-side javascript feature injection-based design for JSP web applications against XSS vulnerabilities. In Cyber Security (pp. 459-465). Springer, Singapore.1