Low Rate Multi-vector DDoS Attack Detection Using Information Gain Based Feature Selection

R. R. Rejimol Robinson and Ciza Thomas

Abstract The number of connected devices is exponentially growing in the world today and they need to work without having any interruption. This scenario is very challenging to cybersecurity and needs proper attention of network administrators, service providers, and users. Implementing security frameworks in this scenario is very difficult because attackers are using very sophisticated easy to operate weapons to launch huge attacks such as Distributed Denial of Service. Intelligently detecting and mitigating the attacks in the network requires the use of machine learning algorithms. This work proposes a strategic way involving feature selection based machine learning for the detection of stealthy attacks. The detection system works by performing ınformation gain-based feature selection as a preprocessing step. This ensures case-based preprocessing of each attack vector present in the traffic and is proved to be effective empirically. The proposed method has been tested using two supervised machine learning classification algorithms, namely Random forest and J48. The evaluation results show that the Random forest algorithm gives a satisfactory True Positive rate of 99.6% in detecting stealthy layer 7 attacks. The overall accuracy obtained is 99.81%. This approach causes the algorithms to exhibit improved performance while doing classification.

Keywords Machine learning · Feature selection · Low rate attacks · Information gain · Stealthy attacks · Network security

1 Introduction

The digitally connected modern world demands uninterrupted connections, even the disruptions are unavoidable. Distributed Denial of Service (DDoS) attack is one such

R. R. Rejimol Robinson (✉)
SCT College of Engineering, Thiruvananthapuram, India
e-mail: rejibz@sctce.ac.in

C. Thomas
Directorate of Technical Education, Thiruvananthapuram, Kerala, India

A. Pasumpon Pandian et al. (eds.), *Computer Networks, Big Data and IoT*, Lecture Notes on Data Engineering and Communications Technologies 66,
https://doi.org/10.1007/978-981-16-0965-7_53

annoyance that makes an online service unavailable. Its impact is devastating unless it is detected and mitigated properly. Formally, define the DDoS as a cyber-attack launched with overwhelming traffic from multiple sources to make the target machine (server) or other network resources unavailable to its intended users temporarily or indefinitely. So it is impossible to stop the attack by simply blocking a single source of the attack. It is the responsibility of network administrator to monitor and supervise their network and guarantee the proper functioning of the network.

There are several strategies for launching an attack and the most prevalent used in these days are zomby attacks or otherwise known as botnet attack. The attack is carried out by a handler that infects vulnerable hosts and recruits them for their purpose. These machines are called zombies on the ınternet and under the control of the handler, zombies are directed to launch an attack by bombarding false packets towards a target to limit its performance or crash it.

Instead of depending on traditional methods of DDoS detection based on Firewalls and Intrusion Detection System (IDS), it is desirable to switch on to machine learning and deep learning based detection strategies. The detection and mitigation system needs to be more agile and intelligent as the attacks are getting more and more sophisticated. But the performance of algorithms is affected due to massive and high-dimensional data received in real-time. Worldwide Infrastructure Security Report by Netscout declared that hackers entered into the era of terabit attacks. According to them, 1.7 TBPS DDoS was recorded in the year 2020 [1]. It indicates the fact that DDoS attack has continued to increase both in size and sophistication.

The stealthiness of DDoS attacks is also getting higher such that multi-vector, slow, and low rate attacks are common. Combination of volumetric and protocol attacks aiming at different layers of network such as layer 7, transport, and network layer are very common nowadays. Multi-vector attacks are also known as polymorphic cyber attacks. They use two or more methods of infiltration. They are launched for a variety of reasons. Some attacks aim to steel sensitive data which is later handed over to a third party merely for monetary benefit and some attacks to take down the network. These are automated attacks and they dynamically change parameters and vectors in response to the defense mechanisms. The Corero Network Security Inc., explained in their blog that sometimes the multi-vector attacks layer different vector types and sometimes they vary the attack vector itself to evade detection. When it is continually modified, it becomes much more difficult to mitigate them. They often start with one vector, such as a simple UDP flood and, if unsuccessful, they try a second technique such as a DNS flood [2]. Hence it is very difficult to detect modern DDoS attacks due to their varying features and different points of entry to infiltrate the network. In this scenario, it is highly essential to do some intelligent preprocessing steps before the machine learning algorithms are applied for detection.

Network traffic data processing comes in two flavors namely, packet-based and flow-based approaches. A flow can be formally defined as a unidirectional stream of IP packets that share a set of common properties; the IP-protocol, source, destination IP addresses, source, and destination ports used. It is often desirable to analyze data inflows rather than packets since it greatly reduces the complexity of data. It is logical

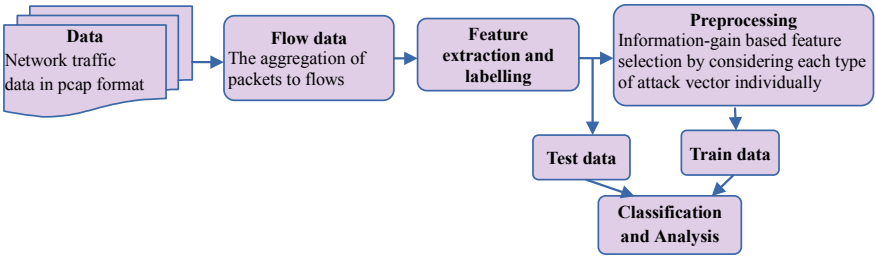

Fig. 1 Workflow diagram of the proposed method

to consider a connection as benign or anomalous rather than packets and hence flow-based analysis is the efficient and fast way of detection. Hence network traffic in the form of flows is used in the proposed work.

The proposed work, mainly concentrating on modern stealthier attacks especially multi-vector stealthy attacks. The workflow diagram showing the conceptualization of the proposed work is given in Fig. 1. The number of benign packets passing through the network per unit time at a particular point is enormous compared to several attack instances. So DDoS attack detection is regarded as an imbalanced dataset problem. Hence it is required to make the features more and more bright to make detection easier. So feature selection is considered an essential preprocessing step in this work. According to Jain et al, feature selection is an effective way of dimensionality reduction and can reduce the complexity of attack detection and thereby increase the detection accuracy [3]. To employ an intelligent strategy initially select the features, and the stealthier attacks. Stealthiness is considered as the resemblance of attack vectors to benign traffic. Different attacking vectors have different prominent features. Hence case-based feature selection is proposed, where each attack vector is treated separately to select features. Making features very prominent is crucial, as it is an imbalanced dataset problem. Feature selection algorithm which can project the maximum information regarding minority attack samples is recommended. So ınformation gain-based feature selection is proposed for doing preprocessing. It is one of the popular feature selection methods due to its computational efficiency and simplicity. This method selects a feature that has high relevance to the output class and also has the highest occurrence rather than simply performing the dimension reduction.

There are two algorithms selected namely J48 and Random forest for doing the performance analysis. The parameters of the algorithms are set in their default settings and it is proposed to evaluate the algorithms with and without doing the information gain-based feature selection on the training data. The important factor of the proposed method is that the classification is being done using a supplied test set rather than doing cross-validation.

The contributions of this work involve:

1. Stealthier multi-vector attack detection with Information gain-based feature selection to get an optimal set of features based on its distinguishing property related to each attacking vector.
2. Enhancing the performance of machine learning algorithms in detecting stealthier attacks.
3. Comparison of methodology with results obtained in other related literature.

1.1 Stealthy Attack Variants

(a) *Low and slow attacks*

These types of attacks use a low volume of data and operate very slowly. Designed to send small amounts of data across multiple connections to keep ports on a targeted server open as long as possible, these tools continue to utilize server resources until a targeted server is unable to maintain additional connections. Uniquely, low and slow attacks may be effective even when not using a distributed system such as a botnet and are commonly used by a single machine.

(b) *Slowloris*

Apart from being a slow-moving primate, Slowloris is an application designed to investigate a low and slow attack on a targeted server. The elegance of Slowloris is the limited amount of resources it needs to consume in order to create a damaging effect.

(c) *TCP flooding*

The attacker exploits the three-way handshake of TCP communication. The attacker sends SYN packets to the server pretending to establish a TCP connection, the server sends SYN-ACK packet back to the client and keeps a port open to receive ACK from the client, but the attacker never sends a final ACK to the server. So the attacker keeps on sending the SYN packets to the server and the server keeps opening a port temporarily for a specific time. The server stops working and responding to legitimate clients after all the ports are utilized.

(d) *UDP attack*

In a UDP attack, the attacker sends a bogus UDP packets to the server using a random port. The server is actually looking for the application on that port. If the service was not running on that port, the server replies with ICMP unreachable message to the attacker. The attacker continuously sends UDP packets and the server also replying ICMP unreachable message back to clients, that will lead to maximum resource consumption of the victim and the network as well. Eventually, the server can not respond to its legitimate user.

The rest of the sections in this article is organized as follows. Section 2 deals with the detailing of the literature survey. Section 3 describes the proposed case-based

feature selection methodology. Section 4 analyze the results and conclusion of the work is presented in Sect. 5.

2 Literature Survey

Due to the proliferation of data to be handled in this new digital world and the high dimensionality of collected data, feature selection is considered as one of the most important factors in determining the efficiency of detection systems. But to have an efficient model, the redundant and less sensitive features need to be dropped.

Combinations of feature selection methods with other optimization techniques are employed in almost all the application domains to improve the detection accuracy of machine learning algorithms. The work of Chuang et al, achieved comparable accuracy when K-nearest neighbor (KNN) with the leave-one-out cross-validation (LOOCV) is used for classifying eleven different gene expression data set. The preprocessing technique employed to bring out this performance is by doing hybrid feature selection methods involving correlation-based feature selection (CFS) and the Taguchi-genetic algorithm (TGA) [4].

According to the work of Gunal et al., a hybrid of both filter and wrapper feature selection steps is being proposed to analyze the redundancy or relevancy of the text features. The experiment done in this work proved the effectiveness of selecting features using different methods rather than stick on to a single method. The combination of the features selected has a profound impact on text classification [5].

The work of Wang et al, mainly deals with DDoS detection based on feature selection involving SU genetic algorithm. The number of features of the NSL-KDD dataset, reduced to 17 from 41 and the machine learning algorithms such as J48 and Random Forest yields 99.8% of detection accuracy. [6] Singh et al, in their work, considers Naive Bayes as the classifier and use information gain-based feature selection and attained 99.5% accuracy in detecting DDoS traffic present in CAIDA 2007 dataset. They have analyzed the results on packet-based data and the features selected are SYN value, ACK value, and Time To Live (TTL) [7].

According to the work of Osanaiye et al., the Ensemble-based Multi-Filter Feature Selection (EMFFS) method is the amalgamation of Information Gain (IG), Gain Ratio, Chi-squared, and relief is used to select important features. The experiments are done on the NSL-KDD dataset and the accuracy is found to be 99.6% using the J48 algorithm. The number of features employed in this method are 13 [8].

In the work of Kamarudin et al, a hybrid feature selection model combines the strengths of the filter and the wrapper feature selection procedure. This hybrid solution selects the optimal set of features in detecting attacks. Correlation feature selection (CFS) together with three different search techniques known as best-first, greedy step-wise, and genetic algorithm are used. The wrapper-based subset evaluation uses a random forest classifier to evaluate each of the features that were first selected

by the filter method. Tested on KDD99 and DARPA 1999 dataset with ten-fold cross-validation in a supervised environment and yields satisfactory results [9].

Lima et al., proposed a smart detection method using machine learning, and this work is designed to detect both high and low volume DDoS attacks. The preprocessing method such as Recursive Feature Elimination with Cross-Validation (RFECV) is used in this work. The datasets CIC-DoS, CICIDS 2017, CSE-CIC-IDS2018, and the ISCXIDS2012 Dataset are mainly employed in this work to evaluate the model [10].

Gu et al., proposed a semi-supervised weighted k-means detection method. A Hadoop-based hybrid feature selection method is used to find the most effective feature set. Then a semi-supervised weighted k-means method using hybrid feature selection algorithm (SKM-HFS) is employed to achieve better performance. The datasets used are DARPA DDoS dataset, CAIDA DDoS attack 2007 dataset, CICIDS DDoS attack 2017 dataset, and real-world dataset [11].

In the work of Wu et al., employed a hybrid feature selection method to detect network anomalies [12]. Wang et al., proposed a combination of sequential feature selection with dynamic MLP to select the optimal features during the training phase and designed a feedback mechanism to reconstruct the detector when perceiving considerable detection errors. The work is mainly done on the NSL-KDD dataset, ISOT, and ISCX [13].

3 Proposed Methodology

During the network traffic data in PCAP format utilize the UDP, ICMP, and TCP flooding attack packet which traces the CAIDA dataset to form multi-vector attack. TC Preplay tool is mainly employed to generate such traffic. TCPreplay is a free Open Source utility suite for editing and replaying previously captured network traffic. It is designed to replay malicious traffic patterns to the Intrusion Detection/Prevention System. All the remaining experiments are done using Python Pandas and WEKA. Pandas is a software library written for the Python programming language for data manipulation and analysis. WEKA, a data mining software, and an open-source software facilitates data preprocessing and implementation of several machine learning algorithms [17].

There are two important phases, one of which is the strategic feature selection phase and the other one deals with the evaluation of machine learning algorithms. Before going to the actual preprocessing step, the data in PCAP format must be processed to form flows and along with that, features are also extracted. Data is the real matter of concern while evaluating the method. CAIDA (Center for Applied Internet Data Analysis) DDoS Attack 2007 dataset is selected to form the sufficient dataset for multi-vector attacks. These traces consist of TCP, UDP, and ICMP flooding instances [15]. To analyze low and slow attacks, Canadian Institute of cybersecurity, CICIDS 2017 dataset has been selected. This dataset contains realistic background traffic and up-to-date attacks [16]. CAIDA data comes in PCAP format only. To implement a

feature extraction module the above process is handled. It extracts features of each flow related to the packet trace namely: Average_Packet_Size, Number_of_Packets, Time_Interval_Variance, Packet size_Variance, Number_of_Bytes, Packet_Rate, and Bit_rate are the seven extracted features [18]. The feature vector formed in this process is represented as $X = (x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7)$, where x_i is the ith feature extracted from the traffic flow, x_7 is label and X represents the feature vector by the feature extraction module. The CICIDS 2017 dataset comes in CSV format, a collection of flow instances with a total of 79 attributes including the label.

The prepared data is represented in CSV format and is ready to give as input to the next phase. The steps involved in the case-based feature selection using information gain is depicted in Fig. 2. The supervised learning algorithms are evaluated to prepare training and test dataset and they are set as 70% of train and 30% test. Then the training data is split into subsets of training data which comprises any one kind of attack vector and normal data. Then the CfsSubsetEval and information gain based feature selection methods are applied to these subsets. The features at the top positions in the rank list produced by these algorithms are selected.

For having a comparison with other feature selection methods, employ the CfsSubsetEval algorithm available in WEKA. This is a correlation-based feature selection method and is based on the logic such that, relevant feature subsets contain features

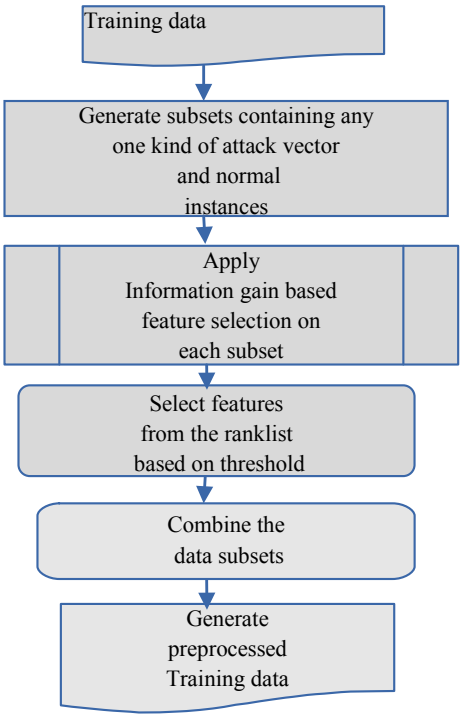

Fig. 2 Flowchart of preprocessing steps involved

highly correlated with the classification, but uncorrelated to each other. The Information Gain (IG) based feature selection is suitable for multi-vector attack feature detection, as it can select the features which contain more relevance which in turn makes the detection easier. It quantifies information based on the contribution of the presence or absence of a feature in making the correct classification decision for any class and is computed according to Eq. (1). The higher the value of mutual information between classes C and feature f, the higher the relevance of feature f in classes C.

$$IG(C, f) = H(C) - H(C|f) \tag{1}$$

where, $H(C) = -\sum_{c \in C} p(c) \log p(c)$ the entropy of the class C and $H(C|f) = -\sum_{c \in C} p(c|f) \log p(c|f))$ is the conditional entropy of class given feature f. The minimum value of $IG(C, f)$ means that $H(C|f) = 1$. That is class C and feature f are not at all related. That is the most distinguishing feature will be the one that is related to a particular class. Due to the feature interactions among the instances, it is not possible to have the ideal case [19].

The next step is to build the new training and test dataset with these selected features. A threshold is set such that those features having information gain value greater than the threshold value are selected as best features. Then new train data is created by joining the subsets of data and it is shuffled to evenly distribute the instances. The next phase of the proposed methodology is the evaluation of machine learning algorithms. J48, the WEKA implementation of decision tree and Random forest algorithms are selected for evaluation. Our earlier work in which the ranking of ten machine learning algorithms has been done. J48 and Random Forest are the algorithms placed in the topmost position in detecting DDoS attacks [20].

A Decision Tree (DT) is one of the well-known supervised classification algorithms in which a tree is generated which acts as a multistage decision system. The concept is based on the measure of the variance of data which demonstrates the presence of different categories of data. A feature vector is assigned with a class label through a sequence of Yes/No decisions along a path of nodes of a DT. The most important factor of splitting criteria for a particular node is to decrease the entropy such that homogeneous vectors can be brought under one particular node since entropy is the measure of impurity [21].

Random forest is an ensemble classification method that combines a collection of classifiers (i.e., decision trees) to make a "forest". Each of the decision trees is generated by using a random selection of attributes at each node to determine the split [22].

4 Result and Discussion

The machine learning algorithms are evaluated based on the True Positive rate (TP) of detecting attack vectors. Among the variety of evaluation metrics such as precision,

recall, F-measure, etc, TP rate is selected to deal with highly skewed data of network traffic. The true positive rate or sensitivity is calculated as given in Eq. (2). TPR is the probability that an actual positive will test positive. TP is the number of instances predicted as positive and TP+FN is the total number of positive samples present in the dataset, where FN is the positive samples misclassified. Misclassification of attacks is costlier than misclassifying benign traffic, so more importance is given to the TP rate of attacks rather than normal traffic. The TP rate comparison is given in Tables 1 and 2.

$$\text{TP\_rate} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

But when evaluating the J48 and Random Forest with this dataset, the detection rate of SlowHTTPtest is considerably very low compared to other attack vectors. The proportionality of the Hulk attack is very high compared to other attacks in the dataset. 33.3% Hulk attack is there while only 0.7% SlowHTTPtest, 0.8% slowloris, and 1.48% GoldenEye in the dataset. The considerable distributional overlap is there between SlowHTTPtest and slowloris attacks. So the TP rate obtained for SlowHTTPtest is very low compared to other attack variants. The features of the GoldenEye attack are very prominent, so the TP rate obtained for this attack is competitively good without any preprocessing. The proportionality of attacks in the CAIDA 2007 dataset is also very low and is 2% only.

The result of the preprocessing step is very important to be analyzed. The CICIDS dataset comes with a total of 79 features and CAIDA 2007 dataset contains 8 features.

Table 1 Table showing the results of CAIDA 2007 dataset

TP rate						
Predicted	CFS		Information gain		Without preprocessing	
	J48	Random forest	J48	Random forest	J48	Random forest
Normal	0.98	0.994	0.993	0.996	0.981	0.992
Attack	0.943	0.989	0.965	0.995	0.995	0.99

Table 2 Table showing the results of CICIDS 2017 dataset

TP rate						
Predicted	CFS		Information gain		Without preprocessing	
	J48	Random forest	J48	Random forest	J48	Random forest
Benign	0.999	0.999	1	1	0.995	0.996
Slowloris	0.991	0.993	0.994	0.994	0.991	0.993
Slowhttptest	0.875	0.874	0.985	0.988	0.856	0.865
Hulk	1	1	1	1	1	1
Goldeneye	0.989	0.989	0.996	0.997	0.989	0.989

Table 3 Comparison of accuracy obtained

Dataset algorithm	CAIDA2007	CICIDS 2017 (%)
Random forest	99.56%	99.81
J48	98.21%	99.77

The number of features selected is greatly reduced especially in the case of CICIDS 2017 dataset. CFSsubsetEval algorithm selects only 6 features for CICIDS 2017 dataset as a whole namely: Destination_Port, Total Length_of_Packet, Bwd_Packets, Init_Win_bytes_forward, Init_Win_bytes backward, and Idle_Max. Similarly, only two features are selected for CAIDA 2007 dataset namely, Time_Interval_Variance and Number_of_Packets. Competitively the number of selected features is more in the case of case-based feature selection. The prominent features of each attack vector are selected separately. It makes each attack more distinguishable even though their proportionality is very low. The bottom line is that a combination of the features selected by information gain based feature selection methods is more effective than the features selected by the CfsSubsetEval. The performance of the machine learning algorithm depends on so many factors among which feature selection is very important. The other important factor is data proportionality. The results show a considerable improvement in detection rate even without any oversampling or synthetic sampling steps, which are the normal preprocessing steps for an imbalanced dataset problem. Effective discriminating power is increased when all the individually selected features together were concatenated.

For having a comparison with the literature published very recently, compute the detection accuracy which is given in Table 3. The literature selected for comparison is briefed in Table 4. The literature which explains the work on modern stealthy attack datasets such as CICIDS 2017 and CAIDA 2007 were considered. This comparison

Table 4 Summary of literature used for comparison

Reference	Feature selection	Model	Dataset	Accuracy (%)
Singh et al. [7]	Information gain	Naive Bayes	CAIDA 2007 and CAIDA anonymous trace 2015	99.5
Lima et al. [10]	NA	Random forest	CIC-DoS, CICIDS2017 and CSE-CIC-IDS2018,	98.6
Gu et al. [11]	Hadoop based hybrid feature selection	Semi-supervised K-means algorithm	DARPA DDoS dataset, CAIDA DDoS attack 2007, CICIDS DDoS attack 2017	99
Singh et al. [14]	NA	Multi layer perceptron with a genetic algorithm	CAIDA 2007	98.04

shows improved performance of machine learning algorithms can be achieved using the proposed method.

5 Conclusion

This work aims to detect stealthier DDoS attacks such as low rate Layer 7 attacks and multi-vector attacks present in the network traffic. Hence it is proposed to have an intelligent approach by preprocessing each attack vector separately to select the most bright features using Information gain. The method is tested with UDP, TCP, ICMP protocol attacks of the CAIDA dataset and the layer 7 low rate attacks such as slowloris, Slowhttptest, Hulk, and GoldenEye attacks of CICIDS 2017 datasets. Machine learning algorithms selected are Random forest and J48 and they are evaluated mainly based on the TP rate of detecting attack vectors. Much improvement of learning algorithms can be claimed over the sub-optimal performance exhibited while performing the classification in the highly imbalanced network traffic. The true positive rate of SlowHTTPtest attack, without any preprocessing is 0.875 but it is improved to 0.985 according to the proposed method and an average TP rate of 0.993 is achieved in detecting multi-vector attacks when tested with supplied test data. The proposed method works superior to any other works found in other literature examined especially in detecting low rate and multi-vector attacks. The accuracy obtained for the Random forest algorithm on CICIDS 2017 dataset is 99.81%.

The work can be extended to include hybrid feature selection methods and can be tested with other combinations of machine learning algorithms as well. Selecting the more distinguishing feature or make the features more and more prominent is the preprocessing method required to make the attack detection more effective.

References

1. Distributed denial of service attack threat report by Netscout. https://www.netscout.com/report/
2. Report on modern DDoS attacks. https://www.corero.com/blog/understanding-and-\\stopping-multi-vector-ddos-attacks/
3. Jain A, Zongker D (1997) Feature selection: evaluation, application, and small sample performance. IEEE Trans Pattern Anal Mach Intell 19(2):153–158 (1997)
4. Chuang L-Y, Yang C-H, Wu K-C, Yang C-H (2011) A hybrid feature selection method for DNA micro-array data. Comput Biol Med 41(4):228–237
5. Gunal S (2012) Hybrid feature selection for text classification. Turkish J Electr Eng Comput Sci 20(2):1296–1311
6. Wang C, Yao H, Liu Z (2019) An efficient DDoS detection based on Su-genetic feature selection. In: Cluster Comput 22(1):2505–2515
7. Singh NA, Singh KJ, De T (2016) Distributed denial of service attack detection using Naive Bayes classifier through info gain feature selection. In: Proceedings of the international conference on informatics and analytics, pp 1–9

8. Osanaiye O, Cai H, Choo K-KR, Dehghantanha A, Xu Z, Dlodlo M (2016) Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. EURASIP J Wirel Commun Netw 2016(1):130
9. Kamarudin MH, Maple C, Watson T (2019) Hybrid feature selection technique for intrusion detection system. Int J High Perform Comput Netw 13(2):232–240
10. Lima Filho FSD, Silveira FA, de Medeiros Brito Jr A, Vargas Solar G, Silveira LF (2019) Smart detection: an online approach for DoS/DDoS attack detection using machine learning. Sec Commun Netw
11. Gu Y, Li K, Guo Z, Wang Y (2019) Semi-supervised K-means DDoS detection method using hybrid feature selection algorithm. IEEE Access 7:64351–64365
12. Wu H, Zhang B, Dong S (2015) A hybrid feature selection method for network traffic anomaly detection. J Phys Conf Ser 1395(1):1. IOP Publishing, 2019
13. Wang M, Lu Y, Qin J (2020) A dynamic MLP-based DDoS attack detection method using feature selection and feedback. Comput Sec 88:101645
14. Singh KJ, De T (2017) MlLP-GA based algorithm to detect application layer DDoS attack. J Inform Sec Appl 36:145–153
15. The CAIDA UCSD DDoS attack 2007 dataset. https://www.caida.org/data/passive/ddos-20070804dataset.xml
16. The CİCİDS 2017 dataset. https://www.unb.ca/cic/datasets/ids-2017.html
17. Hall M, Frank E, Holmes G, Pfahringer B, Reute-mann P, Witten IH (2009) The weka data mining software: an update. In: ACM SIGKDD explorations newsletter, vol 11(1), pp 10–18
18. Karimazad R, Faraahi A (2011) An anomaly-based method for DDoS attacks detection using RBF neural networks. In: 2011 international conference on network and electronics engineering, IPCSIT, vol 11
19. Kent JT (1983) Information gain and a general measure of correlation. Biometrika 70(1):163–173
20. Robinson RR, Thomas C (2015) Ranking of machine learning algorithms based on the performance in classifying DDoS attacks. In: 2015 IEEE recent advances in intelligent computational systems (RAICS). IEEE, pp 185–190
21. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106
22. Liaw A, Wiener M et al (2002) Classification and regression by random forest. R news 2(3):18–22